Car−Parrinello MD Simulations for the Na^+−Phenylalanine Complex in Aqueous Solution

Francesca Costanzo* and Raffaele Guido Della Valle

Dipartimento di Chimica Fisica ed Inorganica, Università di Bologna, viale Risorgimento 4, I-40137 Bologna, Italy

Received: February 27, 2008; Revised Manuscript Received: June 17, 2008

Car−Parrinello molecular dynamics (CPMD) calculations are presented for a Na^+(Phe) complex in aqueous solution and for various stable Na^+(Phe) complexes and $Na^+(H_2O)_n$ clusters in the gas phase (with up to six water molecules). The CPMD results are compared to available experimental and ab initio reference data, to DFT results obtained with various combinations of density functionals and basis sets, and to previous classical mechanics MD simulations. The agreement with the reference data in the gas phase validates the CPMD method, showing that it is a valid approach for studying these systems and that it describes correctly the competing Na^+−Phe and Na^+−H_2O interactions. Analysis of MD trajectories reveals that the Na^+(Phe) complex in aqueous solution maintains a stable configuration in which the Na^+ cation hovers above the phenyl ring, at an average distance of 3.85 Å from the ring center, while remaining strongly bound to one of the carboxylic oxygens of Phe. Constrained MD simulations indicate that the free energy barrier opposing dissociation of the complex exceeds 5.5 kcal/mol. We thus confirm that "cation−π" interactions between alcali cations and the π ring, combined with other kinds of interactions, may allow aromatic amino acids to overcome the competition with water in binding a cation.

1. Introduction

The interaction between a cation and an aromatic ring, known as cation−π interaction,[1] is now recognized as one of the strongest noncovalent binding forces in the gas phase.[2] Cation−π interactions are claimed to occur also in aqueous environment,[2−4] where cations are well solvated and a substantial desolvation penalty must be paid to actually bind a cation. Investigations on protein structures suggest that cation−π interactions may take place between amino acids[2,5,6] and that these putative cation−π interaction sites are often found on the surfaces of proteins, exposed to aqueous solvation.[6−8] A review[3] of the few available measurements on the binding energy in aqueous solution appears to confirm that cation−π interactions between suitable pairs of aminoacids have a stabilizing role within proteins and at their surfaces.

To investigate the competition between cation−π interaction and aqueous solvation, we recently[9] performed a molecular dynamics (MD) simulation of the complex between Na^+ and phenylalanine (Phe) in aqueous solution, by using classical two-body forces. We refer the reader to our previous article[9] for an introduction to this particular cation−π model system. Our classical model reproduced quite well the binding energies and the geometric structure of Phe with H_2O and Na^+, but slightly overestimated the sodium−water coordination. The simulation indicated that the Na^+(Phe) complex survives for a significant time in aqueous solution and that the free energy barrier opposing dissociation of the complex is sizable. Of course, these conclusions rest on the reliability of the force field, which in this case was a modified AMBER[10] model.

We have now repeated the simulation of the Na^+(Phe) complex in aqueous solution, this time with Car−Parrinello MD (CPMD).[11,12] In this method the forces acting on the atomic nuclei are computed ab initio by using density functional theory (DFT) to solve the quantum mechanical problem for the electrons. The wave functions of the valence electrons are expanded in a plane waves basis set, while the core electrons are effectively removed from the calculation by replacing them with pseudopotentials. We have chosen methods successfully tested on similar, but simpler, systems[13−16] and extensively validated the methods by calculating the gas-phase structures of several Na^+(Phe) and $Na^+(H_2O)_n$ complexes. Such a careful validation is especially necessary because our work is the first CPMD investigation involving alkali cations and aromatic systems in aqueous solution and because other ab initio calculations for cation−π interactions in water are scarce.[4] In this paper, we mainly focus on features that are important for understanding cation−π interactions in solution, namely the geometry of the Na^+(Phe) complex, the structure of the water solvation shell around the Na^+ cation, the interference between phenylalanine and water in binding the cation, the free energy barrier opposing dissociation of the complex in water, and the distribution of charges on phenylalanine, Na^+, and water. This distribution, which we have analyzed in terms of Wannier function centers (WFCs),,[17−19] is of interest since it provides information on the role of charge transfer and polarization effects, whose importance in cation−π interactions is a matter of debate. [2,20,21]

The article is organized as follows. The various Na^+(Phe) and $Na^+(H_2O)_n$ reference structures are described in section 2. Details on the CPMD methods and on the actual calculations are supplied in section 3. The validity of using CPMD for studying the interactions of Na^+ with phenylalanine and water, the structure of Na^+(Phe) in water, the free energy barrier opposing dissociation of Na^+(Phe), and the charge distribution are all investigated in section 4. A brief discussion concludes the paper.

* Corresponding author. Tel: +39-051-2093710. Fax: +39-051-2093690. E-mail: costanzo@ms.fci.unibo.it.

10.1021/jp801702v CCC: $40.75
Published on Web 09/13/2008

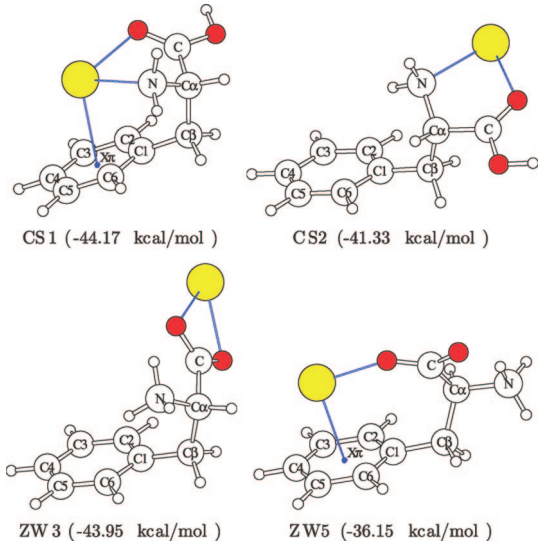

Figure 1. Structures of minimum CPMD energy for the Na^+(Phe) complexes, with IUPAC[28] nomenclature for the carbon atoms. X_π is the center of phenolic ring. Solid lines indicate Na^+−Phe interaction distances reported in Table 2. CPMD binding energies are shown in parentheses. Orientation and labeling match those of refs 23 and 24. Graphics by MOLSCRIPT.[29]

2. Reference Data

2.1. Reference Structures of the Na^+(Phe) Complex. As discussed in the Introduction, we aim at simulating a Na^+(Phe) complex solvated in water. Like all other amino acids, phenylalanine in its free acid form carries simultaneously a carboxylic acid function −COOH and a basic amine function $-NH_2$. In aqueous solution the proton of the −COOH group is transferred onto the amine to give a dipolar zwitterionic form, with $-COO^-$ and $-NH_3^+$ ends.[22] Extensive ab initio calculations[23,24] for the Na^+(Phe) complex in absence of the solvent have identified several stable configurations, in free acid (CS) and zwitterionic (ZW) forms. We have studied the four Na^+(Phe) complexes drawn in Figure 1 and denoted[23,24] as CS1, CS2, ZW3, and ZW5. In the CS1 and CS2 forms the Na^+ cation interacts simultaneously with the carbonyl group C═O and with the amino group NH_2 at the N-terminus. Cation−π interactions between Na^+ and phenyl ring further stabilize the CS1 form,[23] which is the most bound configuration. In the ZW3 form Na^+ interacts with the two oxygens of the carboxilate group $-COO^-$, whereas the ZW5 form presents interactions with one of the two oxygens and cation−π interactions with the phenyl.

2.2. Reference Structures of the $Na^+(H_2O)_n$ Clusters. To verify that CPMD describes correctly the solvation of Na^+ in water, we have compared the computed geometry and energy of several small $Na^+(H_2O)_n$ clusters (drawn in Figure 2) to the available experimental and ab initio results for the gas phase, summarized in Table 3 (subsection 4.2). The experimental data on the binding energies of the $Na^+(H_2O)_n$ clusters[25] are in agreement with the ab initio literature, summarized by Hoyau.[26] Among the available ab initio energies for $Na^+(H_2O)_n$, we have chosen as reference data the large basis sets results by Feller,[27] at the MP2 level (including diffuse functions on the lighter atoms). All structural data also come from these calculations,[27] since no measurements are available.

3. Computational Details

3.1. CPMD Methods. The CPMD calculations for isolated $Na^+(H_2O)_n$ or Na^+(Phe) complexes and for Na^+(Phe) in solution were performed with methods successfully tested in previous investigations on small $Na^+(H_2O)_n$ clusters,[13] on Na^+ in water,[14,16] and on cation−π interactions between amino acids.[15]

The DFT calculations are based on the local density approximation of the exchange-correlation functional augmented by the BLYP generalized gradient corrections, which combines Becke's gradient corrections for exchange[30] with the Lee−Yang−Parr gradient corrections for correlation energy.[31] The more accurate "hybrid" functionals like B3LYP,[32] which include "exact" Hartree−Fock exchange, are unfortunately incompatible with plane waves for performance reasons.[14,33]

Only valence electrons were treated explicitly, and interactions with the ionic cores were described by norm-conserving pseudopotentials generated[16] according to the prescription due to Troullier and Martins.[34] For Na^+ the formal partition in valence and core is ambiguous because ionization eliminates the 3s valence electron. For this reason, the states 2s and 2p in the next inner shell were treated as semicore states, effectively including them in the valence.[16] The Kleinman−Bylander[35] decomposition was used for all the atomic species, with s and p nonlocality (angular momentum dependence) for C, N, O, and Na atoms. H atoms were represented by simple s type pseudopotentials. As usually done for nonperiodic materials, the systems were embedded in a cubic supercell of suitable size, with periodic boundary conditions, and only the Γ point was considered in Brillouin zone integration. The Kohn−Sham orbitals were expanded in plane waves up to an energy cutoff of 70 Ry, which is sufficient to yield converged structural properties for liquid water.[36] Our test calculations on the gas phase complexes indicate that CPMD binding energies are converged within 1.5 kcal/mol.

3.2. Na^+(Phe) and $Na^+(H_2O)_n$ Complexes in the Gas Phase. To check the accuracy of the competing Na^+−phenylalanine and Na^+−water interactions and to validate the CPMD methods, we have calculated the minimum-energy structures in the gas phase of various stable Na^+(Phe) complexes and of $Na^+(H_2O)_n$ clusters with n up to 6. The initial ab initio structures[23,27] were optimized using a quasi-Newton method,[37] combined with direct inversion in the iterative subspace (DIIS).[38]

The difficulties associated with the nonzero charge of the system were tackled by effectively neutralizing the net positive charge with an uniformly charged background and by calculating the long-range Coulomb forces through the associated Poisson's equation, solved with the fast method due to Martyna and Tuckerman's.[39] This procedure requires periodic cell somewhat larger than the system size.[33] In practice, we have found that the cubic cell with sides of 12.1829 Å (≈23 bohr), used for the Na^+(Phe) in liquid water (see subsection 3.3), was also adequate for Na^+(Phe) complexes and for $Na^+(H_2O)_n$ clusters with n up to 4. Sides of 30 bohr were found to be necessary for the larger $Na^+(H_2O)_n$ clusters.

3.3. Na^+(Phe) Complexes in Water. Since solvated amino acids are preferentially in zwitterionic form, and because only the ZW5 conformation presents cation−π interactions, we have chosen this form, among all predicted structures of the complex,[23,24] for our simulation in aqueous solution. The initial configuration for the CPMD simulation was taken from a fully equilibrated classical MD simulation at 300 K, performed in conditions essentially identical to those described in the previous paper,[9] except for a smaller system. In fact, to satisfy constraints on the CPMD computational cost, we choose a cubic cell with sides of 12.1829 Å, containing a Na^+(Phe) complex in its initial ZW5 configuration and 50 water molecules at the experimental density of pure water.[40]

TABLE 1: Binding Energies of the Na^+(Phe) Complexes, Calculated As Differences $E_{Na^+(Phe)} - E_{Na^+} - E_{Phe}$[a]

	B3LYP				BLYP				BLYP CPMD
	6-31G(d)	6-31+G(d, p)	6-311+G(d, p)	6-311+G(3df, 2p)	6-31G(d)	6-31+G(d, p)	6-311+G(d, p)	6-311+G(3df, 2p)	
CS1	−52.81	−47.71	−47.17	−47.74	−51.16	−45.46	−44.95	−45.73	−44.13
CS2	−47.21	−43.91	−43.51	−44.32	−46.24	−42.48	−42.18	−42.97	−41.27
ZW3	−48.13	−44.49	−43.81	−44.18	−47.93	−44.01	−43.48	−43.76	−43.89
ZW5	−41.58	−37.74	−37.09	−37.86	−41.91	−37.60	−36.97	−37.87	−35.95

[a] All binding energies (kcal/mol) are at the B3LYP/6-31G(d) reference geometries.[23] The Gaussian03 energies, with B3LYP and BLYP functionals and the indicated basis sets, include the BSSE corrections.

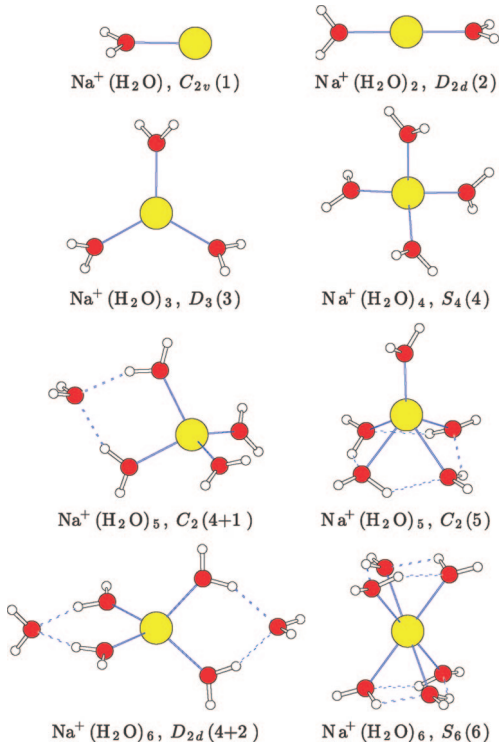

Figure 2. Minimum energy structures of the $Na^+(H_2O)_n$ clusters, labeled by their symmetry group and by their coordination (n_1) or ($n_1 + n_2$) around the Na^+ cation. Here n_1 and n_2 are the number of H_2O molecules in the first and second coordination shells, respectively. Solid and broken lines indicate Na^+−O and H−O bonds, respectively. Orientation and labeling match those of ref 27. Graphics by MOLSCRIPT.[29]

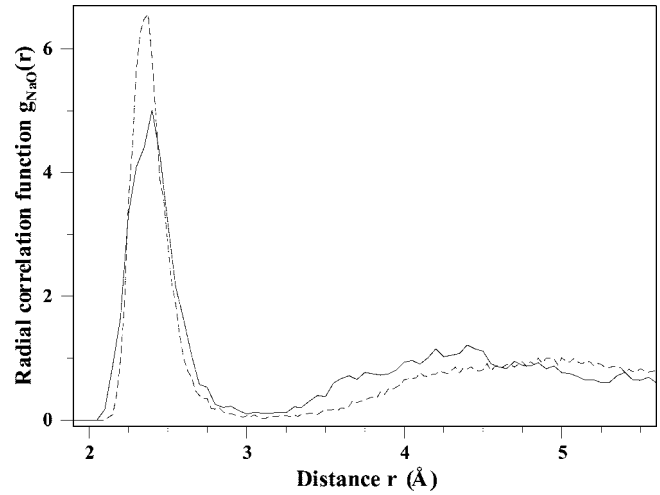

Figure 3. Radial correlation function $g_{NaO}(r)$ between the Na^+ cation and the H_2O oxygens for the Na^+(Phe) complex in water solution. Solid and dashed lines refer to CPMD and classical MD[9] results, respectively.

Hydrogen was substituted by deuterium, to allow for a relatively large time step of 5 au (≈0.12 fs). This choice does not affect the structural properties since the ions are treated as a classical particles in CPMD. A fictitious electronic mass of 800 au was adopted to keep the system on the Born−Oppenheimer surface. To reduce the influence of the initial classical configuration on the CPMD simulation, we performed ≈2 ps of equilibration with a Nosé thermostat[41] at 300 K, followed by ≈10 ps of actual simulation in the NVE ensemble without temperature control.

3.4. Free Energy Profiles. To describe the dissociation of Na^+ from phenylalanine, we have chosen an appropriate coordinate, namely the distance $r = r_{NaC_1}$ between the cation Na^+ and C_1 (the phenolic carbon bearing the aliphatic chain)[28] and then computed the free energy as a function of r. Experimentally, a difference in free energy is determined either from the relative probability of finding the system in a given state or from the reversible work required to transform the system from one initial state into another. Computationally, we have tested both approaches. In the first approach, we calculated the probability distribution $P(r)$, by accumulating a histogram of the distances r during the simulation, and then determined the free energy through the relation[42] $F(r) = -k_B T \ln P(r)$.

In the second approach, we have found the reversible work by using the blue-moon ensemble method,[42] which is formulated in terms of time averages over constrained trajectories where a reaction coordinate r is fixed at a specific value. In our special case in which the reaction coordinate is an interatomic distance, it can be shown[42,43] that the free energy coincides with the potential of mean force (PMF), obtained by integrating the time-averaged force of constraint $\bar{f}$ with respect to r, $\Delta F(r) = \int_{r_0}^{r} \bar{f}(r)\, dr$. Indeed, this special property of the interatomic distances is the reason for choosing the distance from the carbon C_1 as the reaction coordinate, instead of the apparently more appropriate distance from the center X_π of the phenyl ring (see Figure 4). In any case, the two distances have the same meaning and are highly correlated. The PMF describes the free energy profile $\Delta F(r)$ along the reaction coordinate r, with all the other degrees of freedom effectively averaged out.

4. Results and Discussion

4.1. Na^+(Phe) Complexes in the Gas Phase. For a first test of the computational methods, we studied the Na^+(Phe) complex in the gas phase, in the configurations CS1, CS2, ZW3, and ZW5. As reference geometries we selected those obtained by Siu and Ma[23] with the B3LYP functional and 6-31G(d) basis set, since they successfully describe the structure of the ZW5 configuration. We first calculated with Gaussian03[44] the binding energy at the published[23] B3LYP/6-31G(d) reference geometries, using the B3LYP and BLYP functionals with progressively larger basis sets, by including diffuse and polarization functions in the sequence 6-31G(d), 6-31+G(d,p), 6-311+G(d,p), and 6-311+G(3df,2p). Using CPMD (plane waves) we then com-

TABLE 2: Interaction Distances in the Na^+(Phe) Complexes[a]

complex	distance	CPMD	B3LYP/6-31G(*d*)
CS1	r_{NaO}	2.35	2.30
	r_{NaN}	2.53	2.45
	r_{NaX_π}	2.79	2.68
CS2	r_{NaO}	2.29	2.22
	r_{NaN}	2.45	2.38
ZW3	r_{NaO}	2.35	2.26
	r_{NaO}	2.32	2.28
ZW5	r_{NaO}	2.15	2.09
	r_{NaX_π}	2.70	2.58
ZW5 in water	r_{NaO}	2.46 ± 0.13	–
	r_{NaX_π}	3.85 ± 0.39	–

[a] The distances (Å) are for the optimized CPMD structures, for the B3LYP/6-31G(d) reference geometries[23] and from CPMD simulations in solution. All distances are indicated in Figure 1.

puted the binding energies at the reference geometries and, starting from these, we finally determined the minimum energy structures.

Gaussian03 and CPMD binding energies at the reference geometries, calculated as differences $E_{Na^+(Phe)} - E_{Na^+} - E_{Phe}$, are listed in Table 1. The CPMD binding energies at the minima, and the corresponding structures, are instead shown in Figure 1. All E_{Phe} energies are at the geometry in the complex and thus do not include the deformation energy required to bring the uncomplexed neutral molecule into its complexed geometry. The Gaussian03 energies include the correction for the basis set superposition error (BSSE), calculated with the counterpoise procedure.[45,46] Enlarging the basis size and allowing for the Hartree−Fock exchange (included in the B3LYP functional) both decrease the BSSE. For the smallest 6-31G(d) basis set, in fact, the BSSE is ≈5 and ≈6 kcal/mol with the B3LYP and BLYP functionals, respectively. For the largest 6-311+G(3df, 2p) basis set, the BSSE decreases below 1 and 2 kcal/mol with the B3LYP and BLYP functionals, respectively. For both functionals, the largest BSSE is for the zwitterionic configurations.

Since, except for the smallest 6-31G(d) basis set, all Gaussian03 energies appear well converged, the inclusion of diffuse (+) and polarization functions (d,p) is necessary to describe the energetics of the Na^+(Phe) complexes. The Gaussian03 energies (with diffuse functions, BLYP functional, and BSSE corrections) match the CPMD calculations at the reference geometries (which employ the same functional and are free of BSSE). In practice, as it could have been expected, plane waves properly account for the effects of diffuse functions. The good agreement between the two DFT approaches, with the same functional but different basis sets, is an important result which validates the CPMD methods.

The good agreement for the binding energies at the reference structures (Table 1) also applies to the CPMD energies at the minima (indicated in Figure 1), since the energy gains in going to the minima are very small (at most 0.2 kcal/mol). Experimental energies in the gas phase are available only for the most stable configuration CS1, for which a recent review[47] reports experimental binding energies in the range 41.5−49.1 kcal/mol. The good match with our calculations further validates the methods.

We want to stress that even for the configurations CS1 and ZW5, where cation−π interactions are expected to be present, a significant part of the binding is due to other kinds of interactions. In fact, previous ab initio and DFT calculations[48] for the interaction between Na^+ and toluene (which is the aromatic side chain of phenylalanine) gave binding energies in the range 20−22 kcal/mol. Significant interactions with the amino acid backbone must therefore contribute to the much larger binding energies reported in Table 1.

TABLE 3: Binding Energies and Na−O Coordination Distances for the $Na^+(H_2O)_n$ Clusters[a]

system	configuration		CPMD	MP2	expt
$Na^+(H_2O)$	$C_{2v}(1)$	$E_{binding}$	−23.0	−23.3	−24.0
		r_{NaO}	2.26	2.21	
$Na^+(H_2O)_2$	$D_{2d}(2)$	$E_{binding}$	−41.5	−43.1	−43.8
		r_{NaO}	2.20	2.25	
$Na^+(H_2O)_3$	$D_3(3)$	$E_{binding}$	−59.6	−59.5	−59.6
		r_{NaO}	2.29	2.29	
$Na^+(H_2O)_4$	$S_4(4)$	$E_{binding}$	−73.0	−73.9	−73.4
		r_{NaO}	2.34	2.31	
$Na^+(H_2O)_5$	$C_2(4+1)$	$E_{binding}$	−85.1	−87.7	−85.7
		$\bar{r}_{NaO}$	2.34	2.30	
	$C_2(5)$	$E_{binding}$	−81.0	−85.3	
		$\bar{r}_{NaO}$	2.46	2.32	
$Na^+(H_2O)_6$	$D_{2d}(4+2)$	$E_{binding}$	−96.0	−99.5	−96.4
		r_{NaO}	2.32	2.30	
	$S_6(6)$	$E_{binding}$	−87.2	−96.9	
		r_{NaO}	2.41	2.42	

[a] Our CPMD binding energies $E_{binding}$ (kcal/mol) and Na−O coordination distances r_{NaO} (Å) are compared to ab initio values (at the MP2 level)[27] and to experimental energies.[25] Where the distance $\bar{r}_{NaO}$ is not unique, we indicate the average coordination distance $\bar{r}_{NaO}$ in the first shell. Cluster labels are as in Figure 2.

In Table 2 we report various Na^+(Phe) interaction distances in the Na^+(Phe) complexes, determined at the CPMD minimum energy structures shown in Figure 1 and at the B3LYP/6-31G(d) reference structures,[23,24] whose geometries are already well converged. The excellent agreement of the CPMD distances with the reference data indicates that the geometry of the various complexes, like their energy, is well described.

4.2. $Na^+(H_2O)_n$ Clusters in the Gas Phase. As a further test of the methods, we determined with CPMD the minimum energy structures of the $Na^+(H_2O)_n$ clusters with n up to 6. The CPMD structures, drawn in Figure 2, again reproduce correctly the corresponding ab initio configurations.[27] Computed binding energies and Na−O coordination distances are compared in Table 3 to high-quality ab initio results[27] (MP2 level) and experimental[25] energies. We find a good overall agreement which, once again, validates the CPMD results.

The structure of the $Na^+(H_2O)_n$ clusters is driven by the competition between the Na^+−O attraction, which favors high-symmetry structures where all H_2O molecules are coordinated to the Na^+ cation, and the water−water interactions, which favor the formation of a second coordination shell. Direct coordination to the cation, in fact, becomes progressively less favorable when the increasing number of H_2O molecules leads to excessive crowding. In agreement with the ab initio calculations,[27] our CPMD results indicate that the most stable structures for n = 5 or 6 are the arrangements labeled $C_2(4+1)$ or $D_{2d}(4+2)$, which present a first coordination shell with four H_2O molecules, linked by H−O hydrogen bonds to one or two H_2O molecules in a second shell. This is a satisfactory improvement with respect to our previous classical model that overestimated the Na^+−H_2O binding and erroneously favored the arrangements labeled $C_2(5)$ or $S_6(6)$, where all H_2O molecules are in the first shell. [9]

As a matter of fact, the binding energies of these "overcrowded" clusters are now underestimated, especially for $Na^+(H_2O)_6$. This discrepancy is due to a known deficiency of the BLYP functional which, although successfully validated for simulations of liquid water,[36] is not so good for the gas phase, where it tends to underestimate the strength of the hydrogen

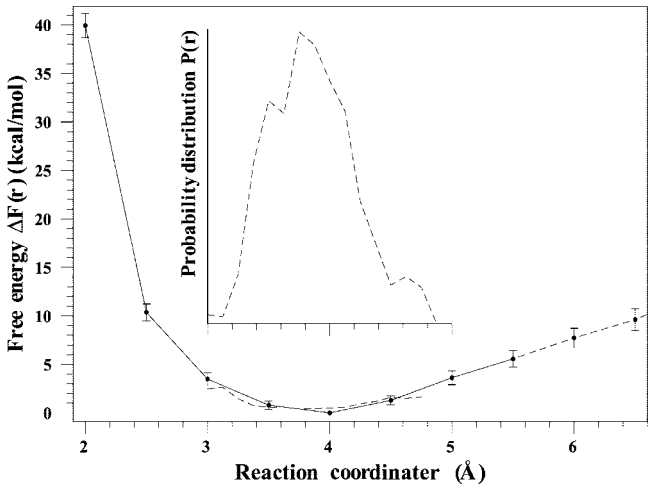

Figure 4. Free energy profile for the dissociation of Na^+(Phe) in aqueous solution. The reaction coordinate is the distance r between Na^+ and C_1 (see Figure 1 for the numbering of carbons). The dashed line for $r \geq r_{fail} = 5.5$ Å suggests unreliable data. The dashed line around $r = 4$ Å indicates the free energy calculated from the probability distribution $P(r)$, shown in the inset. The half-width of the error bars represents the statistical uncertainty on the free energy.

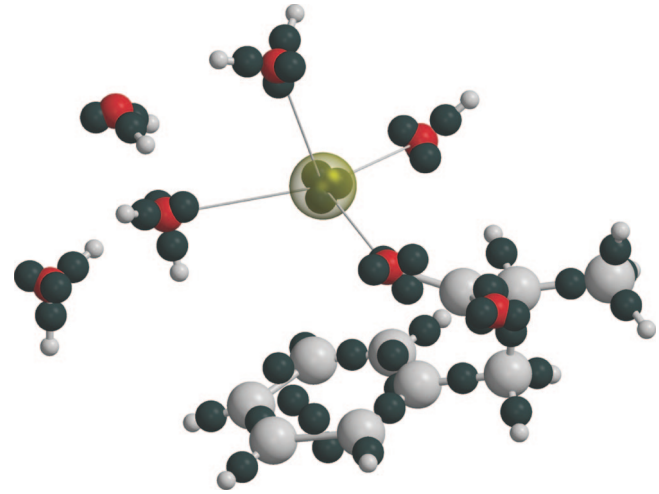

Figure 5. Snapshot of a Na^+(Phe) configuration in water. Five H_2O molecules, three of which coordinated to Na^+, are also shown. The maximally localized Wannier function centers are represented as black spheres. The Na^+ cation is translucent. Graphics by MOLSCRIPT[29] and Raster3D.[55]

bonds.[36,49] As shown in Figure 2, the clusters with an excess of H_2O molecules in the first shell have more hydrogen bonds (four and two bonds for $C_2(5)$ and $C_2(4+1)$, respectively, six and four bonds for $S_6(6)$ and $D_{2d}(4+2)$, respectively). These differences in the number of hydrogen bonds justify the trends of the CPMD energies for the gas-phase clusters. The B3LYP functional, which partially describes van der Waals forces, hydrogen bonding, and other dispersion effects, would evidently be more appropriate for the gas phase clusters.

Since CPMD energies and structures for all gas-phase complexes compare well with the experimental and ab initio literature data on Na^+(Phe) and $Na^+(H_2O)_n$,[23−27,47] we may assert that CPMD describes correctly the Na^+(Phe) binding and its competition with solvation in water.

4.3. Simulated Na^+(Phe) Complex in Water. Once the CPMD method was validated, we simulated the Na^+(Phe) complex in water solution and found that it remained stable in its initial ZW5 configuration during the whole simulation. A snapshot of the simulated system is shown in Figure 5 (subsection 4.5). In Table 2 we compare the average Na^+(Phe) interaction distances in solution to the corresponding distances in the gas phase. The average distances suggest that also in solution the Na^+ cation interacts with the π-ring and with the carboxylic oxygens. Indeed, by analyzing the trajectories as a function of time, we have noticed that the cation hovers above the phenolic ring while remaining bound to one of the two carboxylic oxygens, which occasionally flip and exchange their roles. The r_{NaX_π} distances in solution are larger than in the gas phase, suggesting a weaker binding and a significant solvation effect.

The simplest quantity which reveals the structure of the solvation shell of the Na^+(Phe) complex is the $g_{NaO}(r)$ pair radial correlation function between the Na^+ cation and the oxygens of the water molecules. The $g_{NaO}(r)$ functions computed with our classical model[9] and with the present CPMD approach are shown in Figure 3. The two $g_{NaO}(r)$ functions are quite similar. Both functions present two peaks around 2.40 and 4.75 Å, which correspond to the first and second solvation shells around the Na^+ cation. Since at distances $r \geq 3.5$ Å (second coordination shell and beyond) the two computed $g_{NaO}(r)$ functions are essentially identical and quite featureless, there are no differences in the long-range order. Instead, there are noticeable differences in the first coordination shell, where the peak heights and shapes are different. As previously noticed, our classical model[9] overestimates the Na^+−water bindings and favors shorter average distances $\bar{r}_{NaO}$ and higher coordination numbers $\bar{n}_{NaO}$. For the classical simulation,[9] in fact, we found $\bar{r}_{NaO} = 2.40 \pm 0.01$ Å and $\bar{n}_{NaO} = 4.01 \pm 0.01$, to be compared with $\bar{r}_{NaO} = 2.42 \pm 0.01$ Å and $\bar{n}_{NaO} = 3.29 \pm 0.04$ for the CPMD simulation. These parameters have been estimated by integrating the radial correlation function $g_{NaO}(r)$ up to the first minimum, which for both the classical and the CPMD simulations occurs at $r_{min} = 3.10 \pm 0.10$ Å. The uncertainties on $\bar{r}_{NaO}$ and $\bar{n}_{NaO}$ result from the uncertainty on r_{min}.[9]

The $g_{NaO}(r)$ functions of published CPMD calculations for Na^+ fully solvated in water[16,50] present distances $\bar{r}_{NaO}$ in the range 2.40−2.49 Å and average coordination numbers $\bar{n}_{NaO}$ around 5.2, in agreement with the experimental X-ray data.[51−53] The $\bar{r}_{NaO}$ distances are similar to our CPMD distances for Na^+(Phe) in solution and for the larger $Na^+(H_2O)_n$ clusters in the gas phase (see Table 3). For the coordination numbers $\bar{n}_{NaO}$, the comparison with our results indicates that, in going from fully solvated Na^+[16,50] to a bound Na^+(Phe) complex in water, the Na^+ cation sheds almost two H_2O neighbors. As already mentioned, in fact, the Na^+(Phe) interaction distances (Table 2) show that Na^+ coordinates to the phenyl ring, displacing a single H_2O molecule, and to the $-COO^-$ moiety, displacing another H_2O molecule. This arrangement is visible in the snapshot shown in Figure 5, where only three H_2O molecules find enough coordination space around the Na^+ cation.

4.4. Free Energy Profiles. The computed free energy of the Na^+(Phe) complex in aqueous solution, as a function of the distance $r = r_{NaC_1}$, is shown in Figure 4. We first determined the free energy $F(r)$ from the probability distribution $P(r)$ which, as mentioned in subsection 3.4, was computed by analyzing the distances during the MD trajectory, at essentially zero cost. Though easy to implement, this method provides data only close to the minimum, at energies actually sampled during the simulation. The computed $P(r)$ and the corresponding $F(r)$, both shown in Figure 4, indicate that the Na^+(Phe) complex in aqueous solution presents a rather stable equilibrium configuration with $r \approx 4$ Å.

As discussed in subsection 3.4, we then used the blue-moon ensemble method[42] to compute the free energy at distances r not normally sampled in the simulation. We choose the point at $r = 4$ Å as the zero of the free energy and integrated the average force of constraint $\bar{f}$, separately toward smaller and larger distances, in steps of 0.5 Å. Each constrained simulation involved 0.3 ps of equilibration, followed by data collection for about 2 ps. The statistical uncertainty in the free energy ΔF, indicated in Figure 4, is due to accumulated propagation of errors in the integration and arises from the statistical uncertainty in the mean force $\bar{f}$. This latter uncertainty is given by $\sigma_f/\sqrt{n}$, where σ_f is the standard deviation of the force, which measures the fluctuations about the mean $\bar{f}$, and n is the number of independent samples. Since successive configurations are highly correlated and do not count as independent, we used the procedure discussed by Allen and Tildesley[54] to estimate the effective number n of independent samples in the simulation. With our data, only a configuration every 12 fs may be considered as independent.

The comparison between the free energies calculated with the two procedures, in perfect agreement in the range where they overlap, validates the blue-moon ensemble method. Since the periodic simulation cell has side length $L = 12.1829$ Å (subsection 3.3), the integration cannot be pushed beyond $r = L/2 \approx 6.1$ Å, where the Na^+ cation would be closer to some periodic replica of phenylalanine than to the molecule at the origin, and probably becomes unreliable somewhere around $r_{fail} = 5.5$ Å, where spurious interactions with the replica may begin to be significant. Although this problem does not allow us to estimate the binding energy of the Na^+(Phe) complex, we may still establish that the free energy barrier exceeds $\Delta F(r_{fail}) = 5.5$ kcal/mol. This lower bound is consistent with available ΔF results for the methylammonium−benzene[7] and ammonium−benzene[4] complexes in aqueous media (measured at 5.5 kcal/mol and calculated at 5.75 kcal/mol, respectively). The CPMD barrier is probably larger than predicted by our previous classical MD calculations on Na^+(Phe) in aqueous solution[9] ($\Delta F = 4.5 \pm 1.5$ kcal/mol), although the large statistical error prevents us from making a conclusive statement.

4.5. Wannier Functions. Since the electronic structure is an intrinsic ingredient of CPMD methods, detailed information on the electronic charge distribution may be obtained from the simulation. Following the methods of refs 17−19 which allow the total charge to be partitioned in a chemically transparent and unambiguous way, we have transformed the Kohn−Sham orbitals into maximally localized Wannier functions and then associated the centers of the Wannier functions (WFC) with individual electronic orbitals.

The WFCs of the solvated complex in a typical ZW5 configuration are shown in Figure 5. By examining the WFCs, we have found that the WFCs for Na^+ remain located on the cation, where they appear as the states 2s and 2p. For the water molecules there are two bonding WFCs on the O−H bonds, and two lone-pair WFCs on each oxygen. For phenylalanine there are two lone-pair WFCs on each of the two oxygens of the carboxylate group, and a single bonding WFC on each one the five phenolic C−H bonds. By following the trajectories of the WFCs, it is evident that Na^+ interacts strongly with the lone pair WFCs of the water oxygens. The interaction with the π electrons of the aromatic ring is also evident. The analysis of the charge distribution, in terms of Wannier function centers (WFCs)[17−19] and of Mulliken charge population (data not reported), shows that the charge transfer from the π-ring to the Na^+ cation is negligible. This finding is in agreement with previous gas-phase calculations[21] which indicate that, for the particular case of Na^+, the charge transfer is very small. Such a result, which supports the point of view that many aspects of cation−π interactions can be described by simple electrostatic models,[2,20,21] justifies the remarkable success of our previous classical MD simulation[9] for a Na^+(Phe) complex in aqueous solution. In fact, the comparison with the present CPMD results shows that classical calculations using empirical potential models, although not as accurate as the CPMD calculations, behave quite well in terms of both structure and energetics.

5. Conclusions

Using Car−Parrinello molecular dynamics (CPMD) methods, we have studied a Na^+(Phe) complex in aqueous solution, chosen as a benchmark system to investigate the competition between cation−π interactions and aqueous solvation in binding the Na^+ cation. CPMD results for various stable Na^+(Phe) and $Na^+(H_2O)_n$ complexes in the gas phase, compared to ab initio or experimental reference data, validate the CPMD method for studying the competing Na^+−Phe and Na^+−H_2O interactions and indicate that quantitative accuracy of the calculated properties is to be expected, for both energetics and structural observables.

For the CPMD simulation in solution, among the various stable structures of the Na^+(Phe) complex, we have chosen the so-called ZW5 structure,[23,24] in which phenylalanine is in zwitterionic form and the Na^+ cation interacts with both COO^- and phenyl groups. The ZW5 complex was found to be quite stable in aqueous solution and, in fact, survived for the whole simulation, which spanned more than 10 ps and which allowed for appreciable diffusion and rearrangement in the water solvation shell. Furthermore, constrained MD simulations indicated that the free energy barrier opposing dissociation of the Na^+(Phe) complex is significant. Thus, our work provides strong evidence that the interaction between Na^+ and phenylalanine in aqueous solution, although weaker than in the gas phase, is still sufficient to stabilize the Na^+(Phe) complex. Cation−π interactions with the aromatic side chain of phenylalanine, together with other interactions involving the amino acid backbone, contribute to this stabilization.

Our simulations indicate that the CPMD method is capable of reproducing with acceptable accuracy the characteristics of cation−π systems in aqueous solution. This opens the way toward the study of chemistry involving larger systems (such as peptides) in water.

Acknowledgment. Work done with funds from MIUR (PRIN 2005 and FIRB-RBNE01P4JF through INSTM consortium). We thank Prof. Aldo Brillante, Prof. Michiel Sprik, and Dr. Marialore Sulpizi for helpful discussions, and Prof. Mauro Boero for providing the pseudopotentials.

References and Notes

(1) Dougherty, D. A. *Science* **1996**, *271*, 163–168.
(2) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303–1324.
(3) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250.
(4) Sa, R.; Zhu, W.; Shen, J.; Gong, Z.; Cheng, J.; Chen, K.; Jiang, H. *J. Chem. Phys.* **2006**, *110*, 5094–5098.
(5) Minoux, H.; Chipot, C. *J. Am. Chem. Soc.* **1999**, *121*, 10366–10372.
(6) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9459–9464.
(7) Gallivan, J. P.; Dougherty, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 870–874.
(8) Gromiha, M. M. *Biophys. Chem.* **2003**, *103*, 251–258.
(9) Costanzo, F.; Della Valle, R. G.; Barone, V. *J. Phys. Chem. B* **2005**, *109*, 23016–23023.

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

(11) The CPMD consortium, CPMD code, version 3.7.2.

(12) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

(13) Ramaniah, L. M.; Bernasconi, M.; Parrinello, M. *J. Chem. Phys.* **1998**, *109*, 6839–6843.

(14) Vuilleumier, R.; Sprik, M. *J. Chem. Phys.* **2001**, *115*, 3454–3468.

(15) Sulpizi, M.; Carloni, P. *J. Phys. Chem. B* **2000**, *104*, 10087–10091.

(16) Ikeda, T.; Boero, M.; Terakura, K. *J. Chem. Phys.* **2007126**, 1–9, 034501.

(17) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12487–12865.

(18) Silvestrelli, P. L.; Marzari, N.; Vanderbilt, D.; Parrinello, M. *Solid State Commun.* **1998**, *107*, 7–11.

(19) Silvestrelli, P. L.; Parrinello, M. *Phys. Rev. Lett.* **1999**, *82*, 3308–3311.

(20) Mecozzi, S.; West, A. P., Jr.; Dougherty, D. A. *J. Am. Chem. Soc.* **1996**, *118*, 2307–2308.

(21) Cubero, E.; Luque, F. J.; Orozco, M. *Proc. Natl. Acad. Sci.U.S.A.* **1998**, *95*, 5976–5980.

(22) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*; W. H. Freeman and Co.: New York, 2001.

(23) Siu, F. M.; Ma, N. L.; Tsang, C. W. *J. Am. Chem. Soc.* **2001**, *123*, 3397–3398.

(24) Siu, F. M.; Ma, N. L.; Tsang, C. W. *Chem. Eur. J.* **2004**, *10*, 1966–1976.

(25) Dĭdic, I.; Kebarle, P. *J. Phys. Chem.* **1970**, *74*, 1466–1474.

(26) Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* **1999**, *121*, 8864–8875.

(27) Feller, D.; Glendening, E. D.; Woon, D. E.; Feyereisen, M. W. *J. Chem. Phys.* **1995**, *103*, 3526–3542.

(28) IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature and symbolism for amino acids and peptides. Recommendations 1983. *Eur. J. Biochem.* **1984**, *138*, 9−37.

(29) Kraulis, P. J. J. *Appl. Crystallogr.* **1991**, *24*, 946–950.

(30) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(31) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(32) Becke, A. D. J. *Chem. Phys.* **1993**, *98*, 5648–5652.

(33) Marx, D.; Hutter, J. *Modern Methods Algorithms Quantum Chem.* **2000**, *3*, 329–477.

(34) Troullier, N.; Martins, J. L. *Phys. Rev B* **1991**, *43*, 1993–2006.

(35) Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett.* **1982**, *48*, 1425–1428.

(36) Sprik, M.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1996**, *105*, 1142–1152.

(37) Dennis, J. E., Jr.; More, J. J. *SIAM Rev.* **1977**, *19*, 46–89.

(38) Hutter, J.; Luthi, H. P.; Parrinello, M. *Comput. Mater. Sci.* **1994**, *2*, 244–248.

(39) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810–2821.

(40) *CRC Handbook of Chemistry and Physics*, 84th ed.; CRC Press: Boca Raton, FL, 2003.

(41) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519.

(42) Sprik, M.; Ciccotti, G. *J. Chem. Phys.* **1998**, *109*, 7737–7744.

(43) Meijer, E. J.; Sprik, M. *J. Am. Chem. Soc.* **1998**, *120*, 6345–6355.

(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M.W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.

(45) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024–11031.

(46) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.

(47) Ruan, C.; Rodgers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 14600–14610.

(48) Reddy, A. S.; Sastry, G. N. *J. Phys. Chem. A* **2005**, *109*, 8893–8903.

(49) S. van Erp, T.; Meijer, E. J. *Chem. Phys. Lett.* **2001**, *333*, 290–296.

(50) White, J. A.; Schwegler, E.; Galli, G.; Gygi, F. *J. Chem. Phys.* **2000**, *113*, 4668–4673.

(51) Skipper, N. T.; Neilson, G. W. *J. Phys.: Condens. Matter* **1989**, *1*, 4141–4154.

(52) Caminiti, R.; Licheri, G.; Paschina, G.; Piccaluga, G.; Pinna, G. *J. Chem. Phys.* **1980**, *72*, 4522–4528.

(53) Kameda, Y.; Sugawara, K.; Usuki, T.; Uemura, O. *Bull. Chem. Soc. Jpn.* **1998**, *71*, 2769–2776.

(54) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon. Oxford, UK, 1987.

(55) Merritt, E. A.; Bacon, D. J. *Meth. Enzymol.* **1997**, *277*, 505–524.

JP801702V