Electronic Effects in Biomolecular Simulations: Investigation of the KcsA Potassium Ion Channel

Andrey A. Bliznyuk*,† and Alistair P. Rendell‡

Supercomputer Facility and Department of Computer Science, The Australian National University, Canberra ACT 0200, Australia

Received: March 22, 2004

Ab initio Hartree−Fock (HF) and density functional theory (DFT) calculations have been performed on various model structures containing up to 1019 atoms and derived from molecular dynamics simulations of the KcsA potassium channel. The electrostatic potential and K^+ binding energies are computed and compared to those of Amber molecular mechanics force-field and semiempirical methods. Errors in molecular mechanics results are shown to be predominantly due to neglect of protein atom/protein atom polarization, whereas polarization of the protein by the K^+ ion is a secondary effect. Polarization effects are largest for nearby atoms (3−5 Å) but cannot be neglected even for atoms separated by over 10 Å. Of the semiempirical methods, only the AM1 method gives reasonable results, whereas both PM3 and PM5 failed to describe the K^+ binding energy correctly. Both the HF and DFT methods produce good estimations of K^+ binding energies, but the DFT results suggest abnormally large polarization effects leading to an almost complete loss of charge from the K^+ ion.

Introduction

Today, most biomolecular simulations are performed using relatively primitive models for electronic interactions, e.g., force fields with fixed charge distributions, so electronic effects, such as mutual polarization and charge transfer, are not explicitly included. Rather, in performing most biomolecular simulations, the hope is that electronic effects are small and can be ignored or that they can perhaps be averaged over by modifying the dielectric constant or replacing it with a distance-dependent value.[7] Although there has been some work in recent years to go beyond this limitation, e.g., by using polarizable force fields[1−5] and/or by combining quantum mechanics (QM) and molecular mechanics (MM) methods,[6] these alternatives are not without problems. The most developed among the current polarizable force fields seems to be the SIBFA approach.[1,2] However, even for this highly sophisticated force field, which includes both charge-transfer and polarization energy terms, computations on 100-atom systems produced errors in the range of 3−7%.[1,2] To the best of our knowledge, other force fields are not yet developed sufficiently for applications to large biological molecules.[3−5] Moreover, no QM/MM scheme has yet been developed that permits charge transfer across a QM/MM boundary.

With the advent of linear-scaling QM methods,[8] however, we can now begin to study electronic effects quantitatively and determine exactly how important they are in biomolecular systems. That is, it is now possible to perform a limited number of QM calculations on systems that are of a biologically significant size and then to compare directly the results with equivalent MM calculations.

In this paper, we report on such a comparison for the KcsA potassium channel protein from *Streptomyces lividans*.[9] This channel selectively conducts K^+ ions with high throughput while suppressing conduction of the chemically very similar Na^+ ion by a factor of approximately 10^4. In attempts to understand the operation of this channel at a microscopic level, this system has been the subject of several molecular dynamics (MD) studies.[10−13] However, recent investigations on this and the Gramicidin ion channel[14−16] have indicated that current MM force fields do not accurately model the passage of an ion through an ion channel. As a consequence, comparisons between QM and MM calculations on these systems are of great current interest.

For any simulation purporting to model the operation of an ion channel, it can be argued strongly that the simulation should at least be able to reproduce the electrostatic environment within the ion channel. To this end, we previously reported[17] calculations comparing the electrostatic potential along the center of the KcsA ion channel calculated using a variety of force-field methods with the values obtained using linear-scaling semiempirical theory. This showed that the electrostatic potential could vary by up to 40 kcal/mol depending on whether it was computed using an MM method or semiempirical QM theory. Furthermore, this difference was shown to arise as a result of the neglect of polarization in the MM calculations, and by comparing the electrostatic potentials computed for a variety of small model systems, it was concluded that the semiempirical QM results for the full KcsA ion channel were probably closest to the full ab initio HF results.

This work intends to further our previous study by using a variety of linear-scaling ab initio QM methods, as well as an expanded range of semiempirical QM methods. In part, this new investigation was prompted by a recent publication questioning the validity of semiempirical theory for Coulomb interactions, particularly long-range Coulomb interactions.[18] Noting this, an initial aim has been to verify, using ab initio QM methods, that the conclusions reached in our previous study[17] were correct. A second aim has been to extend our original analysis to include

* Corresponding author: Dr. Andrey A. Bliznyuk. Phone: (61-2) 6125-8924. Fax: (61-2) 6125-8199. E-mail: Andrey.Bliznyuk@anu.edu.au.
† Supercomputer Facility.
‡ Department of Computer Science.

10.1021/jp0487298 CCC: $27.50
Published on Web 08/14/2004

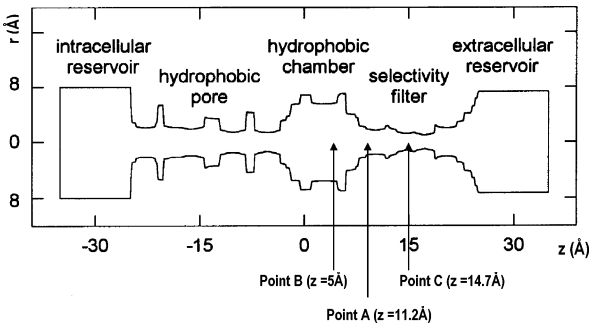

Figure 1. Schematic illustration of the KcsA potassium ion channel showing the location of the three data points considered in this work.

direct evaluation of the binding energy of K^+ inside the channel. Previously, such a direct comparison was not possible because the semiempirical code available to us did not support calculations including K atoms. Since then, however, MOPAC2002[21] has been released, and this restriction has been removed for the AM1,[19] the PM3,[20] and the newly defined PM5[21] methods. Inclusion of the K^+ ion within the channel means that we are now able to gauge the relative importance of ion−protein polarization compared to protein−protein "self"-polarization.

Methods

A schematic illustration of the KcsA ion channel is shown in Figure 1. In this diagram, the center of the pore is aligned with the *z* axis, and the three distinct regions through which the K^+ ion must traverse are marked. The actual structures used in this work are derived from our previous MD simulations[22] as follows:

Structure 1 was derived from MD simulations on the solvated ion channel and without any K^+ ions. All water molecules were subsequently removed to give the coordinate set used here.

Structure 2 was derived from MD simulations on the solvated ion channel and in the presence of a K^+ ion. In these simulations, the K^+ ion was constrained to be located in the hydrophobic chamber at $z = 5$ Å (point B in Figure 1). No water molecules were removed

Structure 3 is similar to structure 2 except that the K^+ ion is now constrained to be located within the selectivity filter at $z = 14.7$ Å (point C in Figure 1). No water molecules were removed.

For structure 1, the electrostatic potential along the central *z* axis was evaluated and compared using a variety of MM and QM methods. As our previous work showed that the maximum disagreement between the MM and semiempirical QM electrostatic potential occurred at $z = 11.2$ Å, we give this location special consideration and have denoted it as point A (see Figure 1). For structures 2 and 3, the electrostatic potential was also computed, but only at points B and C, respectively, and after removal of the K^+ ion. For these structures, the binding energies of the K^+ ion were also evaluated. Although several different sets of coordinates were extracted from the MD simulations, for simplicity and in line with our previous work,[17] the results from only a single set for each structure are presented here.

Structures 2 and 3 were chosen because they represent very different environments for the K^+ ion. In structure 2, the K^+ ion is located in the hydrophobic chamber, the first coordinate sphere consists solely of six water molecules, and the nearest protein atom is approximately 5 Å away. For structure 3, the K^+ ion is located within the selectivity filter and there are three water molecules and several protein atoms (from residues Val76 and Gly77) within a 3-Å radius. Any model that wishes to account for the passage of the K^+ ion through the channel will have to describe accurately both of these very different environments.

TABLE 1: Composition and Number of Atoms and Basis Functions for Shells

shell cutoff	structure 1			structure 2			structure 3		
	atoms	basis	groups[a]	atoms	basis	groups[a,b]	atoms	basis	groups[a,b]
3	104	832	Thr 75	19	137	6x H_2O	150	1192	3× H_2O Val 76 Gly 77
5	196	1588	Val 76 Gly 77	147	1121	12x H_2O Thr 75	206	1668	3x H_2O Thr 75
6	252	2064	Thr 74	292	2178	21x H_2O Ile 100	296	2498	5x H_2O Tyr 78
8	496	4268	Glu 71 Thr 72 Ala 73 Tyr 78	598	4688	27x H_2O Ala 73 Thr 74 Val 76 Phe 103	549	4707	8x H_2O Glu 71 Thr 72 Ala 73 Thr 74 Gly 79
10	848	7120	Gly 79 Met 96 Val 97 Ala 98 Gly 99 Ile 100	1019	8289	31x H_2O Thr 72 Gly 77 Met 96 Gly 99 Thr 101 Ser 102 Gly 104 Thr 107	631	5417	18x H_2O Asp 59
∞	5920	−	−	6551	−	−	6551	−	−

[a] Only residues in the first subunit are listed, but the same residues from other three subunits are also included. Each consecutive shell includes all residues from the previous shell. [b] Total number of water molecules in the shell is shown.

Given the three structures, "shells" of atoms were defined by taking a cutoff radius about the respective points A, B, and C. A shell includes all water molecules and protein residues with any of their constituent atoms within this characterizing distance. For example, shell 5 for structure 2 includes all water molecules and protein residues that have any atoms 5 Å or closer to point B. Once the shell atoms had been defined, any broken peptide bonds were capped using either methylamine (N-Me) or acetyl (ACE) groups. For each of the three structures, shells with radii of 3, 5, 6, 8, and 10 Å were used. The compositions of these shell model structures, in terms of the number of atoms, number of basis functions (when using a 6-31G* basis set), and constituent water molecules and protein residues, are given in Table 1. The smallest system (shell 3, structure 2) contains just six water molecules and the K^+ ion, whereas the largest system (shell 10, structure 2) contains 1019 atoms.

For the various shell models, calculations were performed using the Amber MM force field;[23] the AM1, PM3, and PM5 semiempirical methods; and the ab initio Hartree−Fock (HF) and density functional theory (DFT) methods. The Amber electrostatic potentials and energies were evaluated using an in-house program with standard Amber94 point charges.[23] The semiempirical calculations were performed using MOPAC2002 v.1.1[21] and the parametrized molecular electrostatic potential (PMEP) method therein.[24] The HF and DFT calculations used a 6-31G* basis, which gives reasonable estimates of K^+ binding energies (see, for example, ref 25) and were performed using the Gaussian 03 program.[26] The DFT results used the popular B3LYP energy functional.[27] (For reference, an ab initio HF calculation on the largest system took approximately 10 days to complete using a dedicated 1-GHz HP Alpha processor.)

TABLE 2: Values of the Electrostatic Potential (kcal/mol) at Point A in Various Shell Models Derived from Structure 1[a]

shell cutoff	HF	Amber[b]	AM1[b]
3	−99.3	−118.6 (19.3)	−93.2 (−6.1)
5	−78.4	−95.7 (17.3)	−79.1 (0.7)
6	−86.2	−101.5 (15.3)	−84.8 (−1.4)
8	−108.6	−128.0 (19.4)	−103.3 (−5.5)
10	−98.6	−123.1 (25.4)	−95.5 (−3.3)
∞[c]	–	−146.7	−116.3

[a] See Figure 1 for location of point A. [b] Deviations from HF results are shown in parentheses. [c] Only results for AM1 and Amber calculations are shown, as this system contains over 5900 atoms and is currently too large to be computed at the HF level using our existing computing resources and software.

It should be noted that both the AM1 PMEP method[24] and the Amber point charges[23] were originally derived by fitting to the ab initio HF 6-31G* electrostatic potential data for a range of small molecules. This makes direct comparison of the electrostatic potentials computed by these three methods particularly interesting.

Results

Electrostatic Potential Calculations. In Table 2, we compare the electrostatic potential evaluated at location A for the various shell models based on structure 1. Only the AM1 and Amber methods are considered as these were specifically parametrized to reproduce the HF 6-31G* electrostatic potential (see the Methods section). For Amber and AM1, the values given in parentheses are the differences between the Amber or AM1 values and the corresponding HF values. The results from our previous study[17] are also included in the row labeled ∞, as they correspond to using an infinite shell cutoff or all of the atoms from the parent structure.

As found previously, the value of the electrostatic potential obtained using Amber is always significantly less than the corresponding AM1 or HF value. Comparing the HF and Amber results, we find that, for the 3-Å shell, they differ by 19.3 kcal/mol, and although this reduces to 15.3 kcal/mol for the 6-Å shell, it then grows again to almost 25.4 kcal/mol for the 10-Å shell. In comparison, the AM1/HF differences show much smaller oscillations of between 0.7 and −6.1 kcal/mol. For the largest (10-Å) shell studied, the HF and AM1 results are within 3.3 kcal/mol, suggesting that, for the total system, the HF results would probably be close to the AM1 results. Interestingly, the difference between the AM1 and Amber results for the largest (10-Å) shell is very close to the difference in the complete system (∞ cutoff), i.e., 28.7 compared to 30.4 kcal/mol. This similarity suggests that contributions to the electrostatic potential at point A from atoms that are located farther than 10 Å away are probably relatively well described by Amber point charges. The 10-Å shell is large enough to reproduce more than 80% of the total electrostatic potentials of Amber and AM1 for the complete system. However, the contribution from atoms farther than 10 Å is still very large—more than 20 kcal/mol. This highlights the long-range nature of electrostatic interactions and the difficulty in converging the value of the electrostatic potential by progressively expanding the size of the truncated system.

The electrostatic potential evaluated along the *z* axis of 10-Å shell model from structure 1 is shown in Figure 2. In this system, there are atoms with *z* coordinates ranging from 0 to 23 Å; outside this region, the electrostatic potential would be expected to go smoothly to zero, and this is indeed observed. Within the 0−23-Å segment, there are strong similarities between the shapes of all three curves; however, the Amber electrostatic potential is systematically lower. The differences between the AM1 and HF values are smaller, but nevertheless can be quite significant—up to 10 kcal/mol in some places. Equivalent figures for the smaller shell models show similar trends, in terms of both the close correspondence of the overall shapes and the systematic overestimation by Amber, although the actual values are, of course, different.

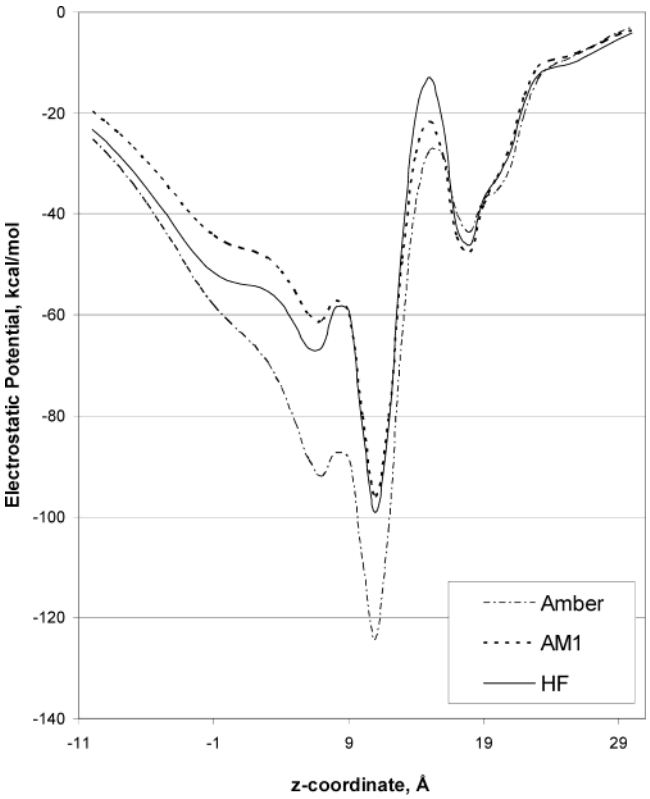

Figure 2. Electrostatic potential evaluated along the *z* axis for structure 1 using the 10-Å shell model

TABLE 3: Values of the Electrostatic Potential at Points B and C in Various Shell Models Derived from Structures 2 and 3, Respectively[a]

shell cutoff	HF	Amber[b]	AM1[b]
	Structure 2 (Point B)		
3	−87.8	−99.0 (−11.2)	−74.2 (13.6)
5	−106.9	−127.3 (−20.4)	−94.0 (12.9)
6	−115.9	−142.4 (−26.5)	−102.5 (13.4)
8	−121.5	−153.4 (−31.9)	−106.6 (14.9)
10	−119.6	−153.7 (−34.1)	−104.0 (15.6)
∞[c]	–	−169.7	−119.1
	Structure 3 (Point C)		
3	−113.7	−145.3 (−31.6)	−103.3 (10.4)
5	−112.5	−141.6 (−29.1)	−102.1 (10.4)
6	−101.5	−129.2 (−27.7)	−94.3 (7.2)
8	−124.6	−157.9 (−33.3)	−112.5 (12.1)
10	−117.3	−148.9 (−31.6)	−106.9 (10.4)
∞[c]	–	−153.9	−110.1

[a] See Figure 1 for locations of points. [b] Deviations from HF results are shown in parentheses. [c] Only Amber and AM1 results are shown, as this system contains 6551 atoms and is currently too large to be computed at the HF level using our existing computing resources and software.

The values of the electrostatic potentials for the various shell models based on structures 2 and 3 with the K^+ ion located at points B and C, respectively, are reported in Table 3. Again, Amber systematically overestimates the electrostatic potential,

deviating from the HF results by more than 30 kcal/mol for the largest shell models. The AM1/HF differences are smaller, at around 10−15 kcal/mol, and of opposite sign. Interestingly, for structure 2, the Amber/HF difference increases gradually from 11.2 to 34.1 kcal/mol as the shell size increases, whereas for structure 3, the difference remains almost constant at around 30 kcal/mol. The AM1/HF differences are almost constant for both structures. As for structure 1 and point A, the difference between the Amber/AM1 results in the 10-Å clusters are very close to the corresponding differences in the complete systems (∞ cutoff).

As each shell includes all atoms from any smaller shell, and as the effects of the capping groups are likely to be small, we can assess the impact of expanding the size of the shell model on the value of the potential by defining the following function

$$\Delta_x^i = [\phi(r)_{\mathrm{HF}}^i - \phi(r)_{\mathrm{HF}}^{i-1}] - [\phi(r)_x^i - \phi(r)_x^{i-1}] \qquad (1)$$

In this equation, the index i represents a particular shell model, and $i - 1$ is the next smaller shell model for the same coordinate set. The electrostatic potential at point r, normally the same location as used to define the shell model, is given by $\phi(r)$, and the subscript x represents either Amber or AM1. If the HF values are considered to be "correct", this function can be viewed as the distance-dependent "error" in the electrostatic potential, e.g., $\Delta_x^{10\ \text{Å}}$ includes the error in the electrostatic potential due to atoms farther than 8 Å from the point of evaluation.

The value of any given Δ_x^i can be consider to arise from two effects: (i) how "new" atoms, i.e., those in shell i but not in shell $i - 1$, interact among themselves and (ii) how new atoms interact with the "old" atoms, i.e., those in both shell i and shell $i - 1$. For example, if we consider lack of polarization to be the primary source of error in the Amber results, one contribution to $\Delta_{\mathrm{Amber}}^i$ would be due to the *new* atoms polarizing other *new* atoms, and another would be due to *new* atoms being polarized by *old* atoms. This sort of error is exactly the same as is present in QM/MM calculations that treat distant atoms using molecular mechanics. For instance, if we assume an HF/Amber QM/MM model with the boundary defined as the interface between the 8- and 10-Å shell models, then the exact electrostatic potential should be $\phi_{\mathrm{HF}}^{10\ \text{Å}}(r)$. The electrostatic potential of our hypothetical QM/MM system is $\phi_{\mathrm{HF}}^{8\ \text{Å}}(r) + [\phi_{\mathrm{Amber}}^{10\ \text{Å}}(r) - \phi_{\mathrm{Amber}}^{8\ \text{Å}}(r)]$. The difference between these two values is exactly $\Delta_{\mathrm{Amber}}^{10\ \text{Å}}$ as defined by eq 1.

The values of Δ_x^i for all three structures with the electrostatic potentials computed at their respective A, B, and C locations are plotted in Figure 3. At first, it might appear that, because of the $1/r$ behavior of the electrostatic potential, this function should approach zero as the shell size progressively expands. However, as the number of atoms increases with the shell size, the overall result is not so simple. For structure 1 using Amber, we find that atoms farther than 8 Å still contribute 5 kcal/mol to the difference between the HF and Amber electrostatic potential. For structure 2, the value goes to zero with increasing distance, but for structure 3, the contribution oscillates. For structure 3, the absolute Amber error is considerably larger than that for AM1, although the amplitudes of the oscillations seen in Figure 3 are of almost equal size. Overall, there appears to be no obvious similarity in the behavior of the error function for these three different structures, and although the error is greatest for the smaller shells and then decreases, in general, it cannot be neglected even at distances larger than 8 Å.

TABLE 4: K^+ Binding Energies for K^+ Ion Located at Points B and C in Various Shell Models Derived from Structures 2 and 3, Respectively[a]

shell cutoff	HF	B3LYP	Amber	AM1	PM3	PM5
Structure 2 (K+ Ion at Point B)						
3	−85.6	−85.8	−92.8	−73.9	−79.6	−81.3
5	−111.5	−111.6	−123.4	−104.1	−92.7	−84.2
6	−123.8	−124.0	−139.2	−115.9	−94.9	−84.3
8	−131.6	−130.9	−150.6	−122.5	−98.5	−86.9
10	−131.0	−131.3	−151.0	−122.4	−98.4	−84.7
∞[b]	–	–	−167.0	−141.2	−117.4	−100.0
Structure 3 (K+ Ion at Point C)						
3	−126.0	−121.8	−141.1	−116.9	−111.7	−120.8
5	−126.0	−121.1	−137.7	−116.4	−107.9	−116.3
6	−116.5	−113.4	−125.7	−108.7	−98.4	−104.8
8	−141.6	−136.5	−154.6	−130.5	−118.7	−123.7
10	−134.6	−130.0	−145.7	−125.4	−113.5	−116.3
∞[b]	–	–	−150.9	−134.7	−120.9	−119.2

[a] See Figure 1 for locations of points. [b] System too large to be calculated using HF or B3LYP methods.

Binding Energy Calculations. The binding energy of the K^+ ion to the channel was evaluated using the following equation

$$E(\mathrm{bind}) = E(\mathrm{protein} + \mathrm{K}^+) - E(\mathrm{K}^+) - E(\mathrm{protein}) \qquad (2)$$

For the HF and DFT calculations, the energies of the isolated K^+ ion and protein were computed using the combined basis of the entire complex, thus taking into account the basis set superposition error. The results for structures 2 and 3 are given in Table 4 (structure 1 did not contain a K^+ ion). As well as including values for Amber, AM1, and HF, we also include data for two additional semiempirical methods, PM3[20] and PM5,[21] and for DFT using the B3LYP[27] functional and a 6-31G* basis set. Inclusion of PM5 is of interest as this method is claimed to be a substantial improvement over AM1 and PM3.[21] As yet, full details of the PM5 method have not been published, nor have the parameters for the K atom with any of these methods, although the PM5 method and parameters for K atoms are now included in the standard released version of MOPAC2002.[21]

In line with the electrostatic potential results, we find that the HF K^+ binding energies (Table 4) converge smoothly for structure 2 as a function of cluster size, whereas they show much larger fluctuations for structure 3. The Amber results are overestimates compared to the HF values, albeit by a smaller margin than for the electrostatic potential, whereas the AM1 values are underestimates. Comparing AM1 to PM3 and PM5, these alternative semiempirical approaches tend to give results that are in worse agreement with the HF values, the only exceptions being for the smallest (3-Å) clusters. For structure 2, the DFT results are very close to the HF values, and although they are further from the HF results for structure 3, the difference is a relatively constant underestimate of about 5 kcal/mol.

Accurate computer modeling of ion conductance through a channel requires the correct description of relative energies only. For this reason, and to highlight further the trends outlined above, in Figure 4, we plot the binding energies obtained for the various methods relative to their computed values in the respective 3-Å shell models. This clearly shows that the Amber results for structure 2 are overestimates, especially for the 8-

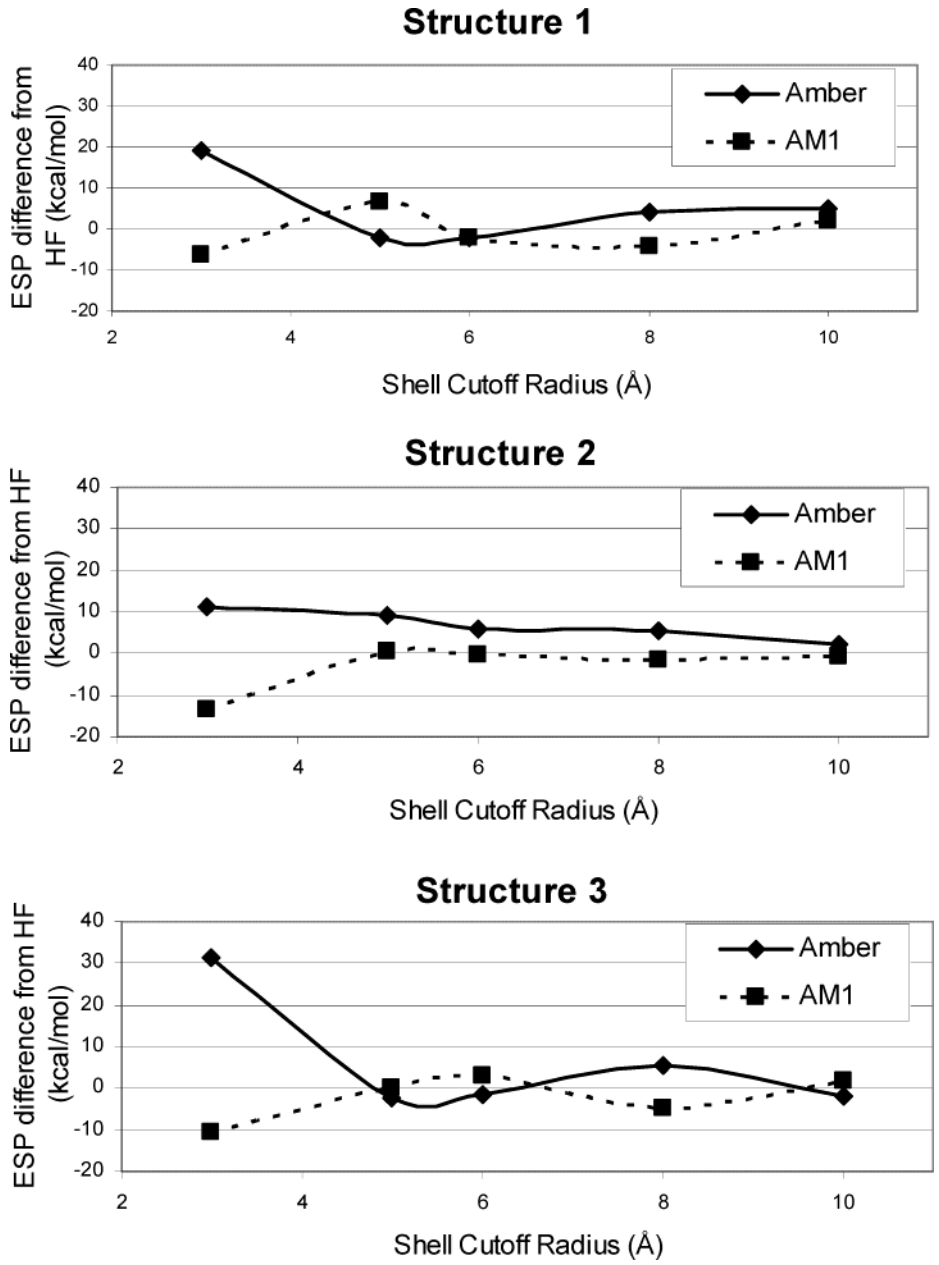

Figure 3. Difference between the values of the Amber and AM1 electrostatic potentials (ESPs) computed at points A, B, and C for structures 1, 2, and 3, respectively, and their various shell models.

and 10-Å shell models, whereas they are actually underestimates for structure 3 by approximately 7 and 3.5 kcal/mol for the 8 and 10-Å shell models, respectively. Of the three semiempirical methods, AM1 clearly performs better, whereas the general tendencies of both the PM3 and PM5 results are incorrect, especially for structure 2. This error is probably due to poor semiempirical parameters for the K^+ ion, although such a result is surprising given that the current MOPAC2002 Web page[21] claims that PM5 gives very good estimations for the heats of formation for K-containing molecules. (The average deviation is claimed to be only 1.8 kcal/mol, compared, for example, to carbon-containing molecules, for which the deviation is 5.7 kcal/mol[21]). This suggests that either a poor choice of test molecules was made in deriving the K parameters or there is another more fundamental problem with the method. Unfortunately, it is impossible to make a definitive conclusion regarding this matter until more details concerning the method have been published. Needless to say, both the PM3 and PM5 methods are unsuitable for energy evaluations on the potassium ion channel and will therefore not be considered further.

In a fashion similar to the electrostatic potential, we define a binding energy error function as

$$\Delta E_x^i = [E(\text{bind})_{\text{HF}}^i - E(\text{bind})_{\text{HF}}^{i-1}] - [E(\text{bind})_x^i - E(\text{bind})_x^{i-1}] \quad (3)$$

For structures 2 and 3 and the Amber, AM1, and B3LYP methods, the values of this function are plotted in Figure 5. For structure 2, Amber systematically overestimates the contributions to the binding energy, and B3LYP produces results that are almost identical to those of HF, whereas the AM1 results oscillate. For structure 3, the function oscillates around 0 for all methods, although with large-magnitude oscillations for Amber. Surprisingly, the results for AM1 and B3LYP are almost identical. For both structures, the binding energy error function is greatest for the smallest shell model and decreases for the

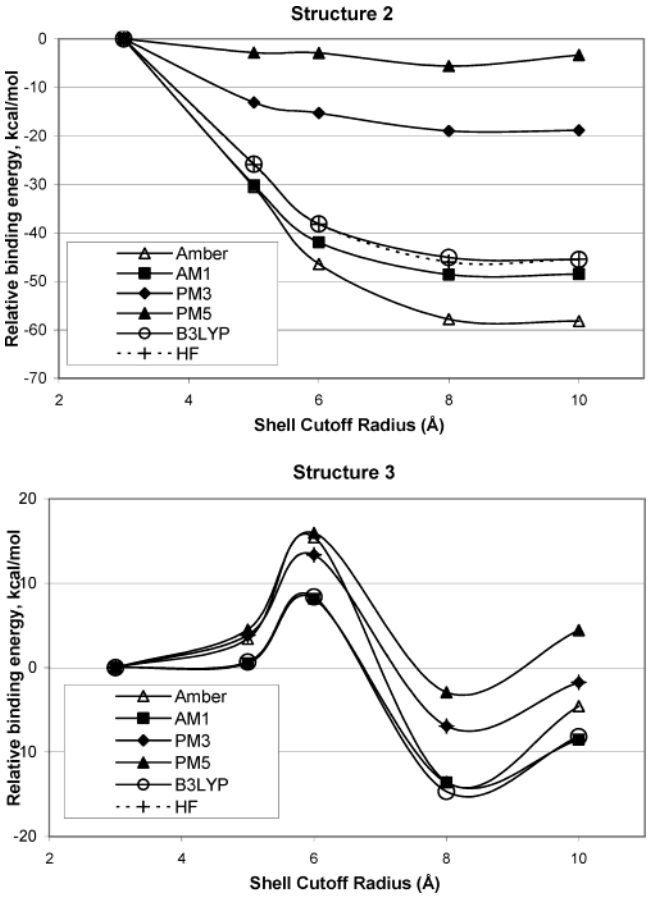

Figure 4. Binding energies for the different shell models relative to the 3-Å shell model for the K^+ ion located at points B (structure 2) and C (structure 3).

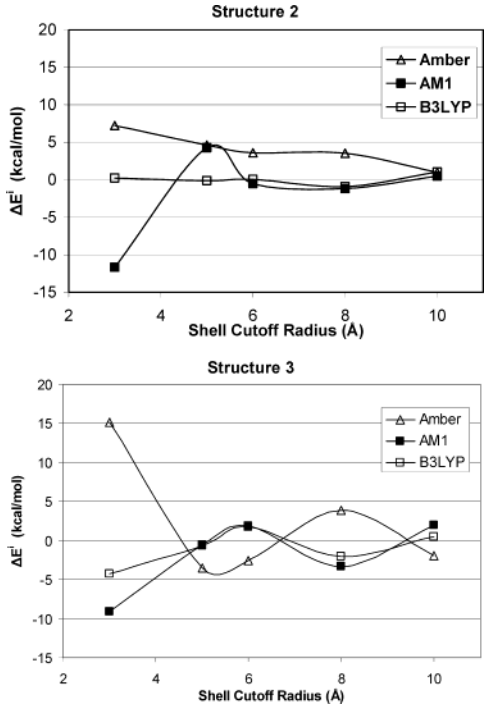

Figure 5. ΔE_x^i (see eq 3) computed using the Amber, AM,1 and B3LYP methods for the binding energy of a K^+ ion located at points B and C in structures 2 and 3, respectively.

larger shell models; however, it is still relatively large for the 10-Å shell models. The only place where AM1 produces a larger error than Amber is for the smallest (3-Å) shell model from structure 2. Because this model consists of just six water molecules and the K^+ ion, this result might indicate that the AM1 description of the K^+/water interaction is not particularly good.

Discussion

One of the challenges for computational studies of the potassium channel is to reproduce the experimental ion current. Our calculations clearly show, however, that direct application of the Amber force field to this problem will give incorrect results due to large and nonsystematic errors in describing both the electrostatic potential and the K^+ binding energy within the channel. For the 10-Å shell model, the errors in electrostatic potential at points A, B, and C are 25.4, 34.1, and 31.6 kcal/mol for structures 1, 2, and 3, respectively (Tables 2 and 3), and the errors in the K^+ binding energies are 20.0 and 11.1 kcal/mol for structures 2 and 3, respectively (Table 4). The variations between the values in each of these two data sets suggest that the Amber results have a relative error of at least 10 kcal/mol. In order for ion permeation to occur, the maximum barrier experienced by the K^+ ion should be close to zero[13] or at most 1−3 kcal/mol.[10] This is essentially the error with which the barrier should be estimated. Unfortunately, from the above results, it appears that the neglect of electronic effects within the Amber force field leads to much larger errors, and on the basis of our previous work,[17] we expect similar errors to be found for all other nonpolarizable force fields.

The above conclusion concurs with results reported recently by two other groups trying to use current MM force fields to study ion channel conductance.[15,16] In the first case, it was necessary to adjust the parameters of the K^+ ion to reproduce better its relative solvation energies in water and *N*-methylacetamide before reasonable results could be obtained from MD simulations.[14,16] Our calculations show, however, that the problem lies not with the K^+ ion parameters but with the failure of current MM force fields to reproduce the electrostatic environment within the protein. Adjusting the K^+ parameters will only introduce some average correction to the protein/K^+ ion interaction, and the results obtained[16] are unlikely to represent the true energetics of K^+ conduction. In the second case, an attempt to evaluate the potential of mean force (PMF) for a K^+ ion inside the gramicidin A channel using the CHARMM and GROMACS force fields failed.[15] This failure was attributed to a lack of polarization in the force fields used, and no attempt was made to change the MM force-field parameters. Our results support this conclusion, although it remains to be seen whether, for example, a polarizable MM force field would be able to solve this problem.

In considering ion channel conductivity, it is often assumed that electronic changes in the protein atoms (often referred to as "polarization effects") resulting from the presence of the ion in the channel are of prime importance. Indeed a recent DFT study by Guidoni and Carloni[28] on the KcsA channel seems to indicate that this is the case. Our calculations, however, suggest otherwise. As shown in Figure 6, errors in the values of the electrostatic potential and K^+ binding energies track each other very closely. Because the electrostatic potential is computed in the absence of the K^+ ion, any error in its value is caused solely by poor representation of protein atom/protein atom interactions. On the other hand, the error in the binding energy of the K^+ ion has two sources: poor representation of protein/protein interactions and poor representation of the protein/K^+ interactions. If the electronic effects due to the K^+ ion were small, the errors in the binding energies would closely follow those

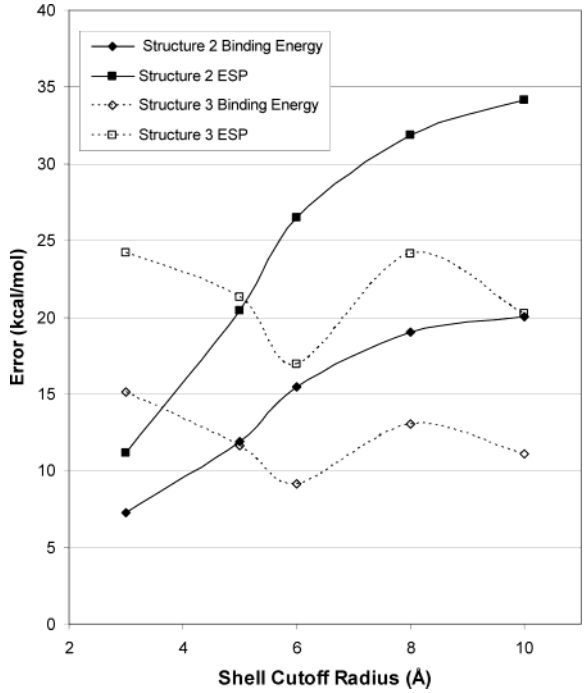

Figure 6. Difference between the HF and Amber values of the electrostatic potential (ESP) and K^+ ion binding energy as a function of shell model size for structure models 2 and 3. Values are correspond to points B and C in structures 2 and 3, respectively.

TABLE 5: Mulliken Charges (au) on K^+ Ion Located at Points B and C in Various Shell Models Derived from Structures 2 and 3, Respectively[a]

	structure 2 (point B)		structure 3 (point C)	
shell cutoff	HF	DFT	HF	DFT
3	0.745	0.580	0.572	0.277
5	0.617	0.369	0.565	0.261
6	0.490	0.169	0.540	0.221
8	0.433	0.079	0.532	0.195
10	0.426	0.068	0.530	0.193

[a] Values calculated using a 6-31G* basis set and the HF and DFT/B3LYP methods.

in the electrostatic potential. If, however, such errors were large, there would be little, if any, similarity between these two results. The fact that we find the error in binding energy to almost exactly follow the error in the electrostatic potential strongly suggests that the primary source of error is poor representation of protein atom/protein atom interactions. The only exception to this observation occurs for structure 2 and the 3-Å shell model, but as this case involves only six water molecules, it is not surprising that the K^+ ion has a large influence.

The main reason Guidoni and Carloni[28] concluded that protein/K^+ electronic effects were dominant appears to be due to the fact that they based their conclusions on electron density evaluations and not on energetics as was done here. It is well-known that DFT methods have a tendency to overemphasize the importance of polarization in large molecules (see, for example ref 29). To illustrate this tendency, we list, in Table 5, the Mulliken charges on the K^+ ion computed by HF and DFT for the various shell models. There is a substantial difference in the charge on the potassium atom depending on which of these two methods is used: the DFT results suggest that electron density from the surrounding atoms migrates to such an extent that the K^+ ion is almost neutral. Clearly, this is incorrect, reinforcing the view that electron density results obtained from DFT calculations should be viewed with extreme caution, especially for large molecules.

In the Introduction, we mentioned that one proposed method to account for electron polarization in MM calculations involves modifying the dielectric constant, either by giving it a non-unit value or by replacing it with a distance-dependent function. In the first case, a value of between 2 and 4 is often proposed to account for a protein's interior (see, for example, ref 12 and the literature cited therein). On the basis of the ratio of the Amber and HF electrostatic potentials, our calculations suggest that a value of $\sim$1.3 would be more appropriate for the potassium channel. In the second case, an S-type dielectric function of the form described in ref 7 is often used. The main feature of this approach is that it reduces electrostatic interactions for nearby atoms only slightly, but greatly decreases those for more distant atoms. Our calculations show that a function that does exactly the opposite is required, i.e., the largest error in electrostatic potential is due to atoms at distances of less than 5 Å from the point of evaluation (Figures 3 and 5) and that this interaction should be reduced most. This is not a surprising result: protein structures put nonconsecutive residues very close to each other, and this gives rise to substantial local distortions in the electron density. Throughout the whole protein, however, the average distortion is approximately constant. Thus, if you are interested in the electrostatic potential close to a specific point, be that in the selectivity filter of the KcsA ion channel or more generally at some active site, electronic distortions close to that location will contribute most to the overall error, whereas those farther away will be reduced by the $1/r$ dependence of the electrostatic potential. This effect is not unique to the KcsA channel and has indeed been observed in recent calculations on mutants of the prion protein.[30]

Why, then, do many MM force-field applications prove to be very successful? The results presented here provide some explanation as to why this is the case. Specifically, most chemical processes require an accurate description of relative properties only. For example, free energy perturbation calculations of local K^+ ion rearrangements[31] are likely to be quite accurate, as the error in electrostatic potential at nearby points will be roughly the same and will therefore cancel. Similarly, PMF calculations on the KcsA channel require that the errors in electrostatic potential be nearly identical both inside and outside the channel (i.e., when the K^+ ion is located in pure water), and our previous calculations showed[17] that, in the models used to date,[17,22] this is probably true. This might also go some way toward explaining why PMF calculations were successful after the K^+ parameters[16] had been adjusted to reproduce the relative solvation energies of the K^+ ion in water and liquid *N*-methylacetamide: solvation of the K^+ ion in *N*-methylacetamide models the inside of the ion channel, and the relative error between having the K^+ ion located inside and outside the ion channel is thus reduced. In contrast, the results in ref 15 suggest that the error difference between a K^+ ion solvated in water and one located within the gramicidin A channel is large, leading to a failure in the PMF method.

Considering QM/MM methods, it appears that use of these schemes will not automatically guarantee correct inclusion of electronic effects. As we have shown, the error in the electrostatic potential obtained from the fixed-charge Amber force field can be quite large, even in the 10-Å shell model. Using QM methods close to the K^+ ion will reduce this absolute error but will not eliminate it. Moreover, care would have to be taken to ensure that the electronic description was improved consistently over all sampled configurations; otherwise, a bias would be

introduced, leading to increases in the relative errors and poorer overall results.

Conclusions

Molecular mechanics force-field calculations and QM calculations using semiempirical, HF, and DFT methods have been performed using various coordinate sets taken from MD simulations on the KcsA potassium channel. Using these methods, the electrostatic potentials and binding energies for a K^+ ion located within the channel have been evaluated and compared. The results show that the main source of error in MM calculations on this system arises from the inability of fixed-charge MM force fields to account correctly for electronic distortions that occur as the protein residues adopt their 3D structure. Not surprisingly, electronic distortions are most important when describing nearby interactions, and their effects decrease with distance, but even at 10 Å, they are not negligible. This work illustrates the importance of the electron density rearrangements due to residues that are close in space. Thus, the methodology of the recently developed approach to protein−ligand energy evaluations[32] in which the energy is computed as a simple sum of interactions of individual residues with the ligand is highly questionable.

Of the three semiempirical methods considered in this work, the AM1 method clearly provides the best description of the KcsA channel: for all but the smallest systems, the AM1/HF difference was found to be considerably smaller than the Amber/HF differences. Surprisingly, the results for PM3 and PM5 were quite poor, and for PM5, even some of the trends were incorrect.

In general, the HF and DFT results were in close agreement. DFT, however, was shown to overestimate dramatically electronic polarization, predicting that a K^+ ion located within the ion channel is almost neutral. This is clearly unphysical and reinforces conclusions from other work[29] that the electron density obtained from DFT calculations on large charged systems should be interpreted with great caution.

In summary, our work shows that existing fixed-charge MM force fields have substantial difficulty in describing accurately the electrostatic environment within protein systems. Polarizable force fields, although offering the possibility of a better electrostatic representation, appear to be some way off in terms of being readily available.[4,5] In the meantime, being able to perform a limited number of ab initio calculations on a large fraction of the atoms contained in an MM simulation is extremely valuable, both permitting researchers to make quantitative estimates of the likely errors within existing force fields and aiding them in the development of new polarizable ones. In addition, large-scale ab initio methods can also assist in the development of new more accurate electrostatic potential methods[33] or new semiempirical methods,[34] and as the speed of these methods continues to increase,[8,35] they can become viable alternatives to MM schemes. It might also be possible to use large-scale QM calculations to reparametrize existing MM force fields for specific proteins and structures. If these were used in subsequent MD simulations where gross geometrical changes are minimized, they might offer a route to providing very accurate results for specific problems.

Acknowledgment. The authors gratefully acknowledge financial and computing infrastructure support from the Australian Partnership for Advanced Computing (APAC). A.P.R. also gratefully acknowledges support from Sun Microsystems and Gaussian Inc.

References and Notes

(1) Gresh, N.; Derreumaux, P. *J. Phys. Chem. B.* **2003**, *107*, 4862−4870.

(2) Antony, J.; Gresh, N.; Olsen, L.; Hemmingsen, L.; Schofiled, C. J.; Bauer, R. *J. Comput. Chem.* **2002**, *23*, 1281−1296.

(3) Tabacchi, G.; Mundy, C. J.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **2002**, *117*, 1416−1433.

(4) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515−1531.

(5) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236−242.

(6) Gao, J. *Rev. Comput. Chem.* **1996**, *7*, 119−185.

(7) Hassan, S. A.; Guarnieri, F.; Mehler, E. L. *J. Phys. Chem. B* **2000**, *104*, 6478−6489.

(8) Goedecker, S.; Scuseria, G. E. *Comput. Sci. Eng.* **2003**, *5*, 14−21.

(9) Doyle, D. A.; Cabral, J. M.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. *Science* **1998**, *280*, 69−77.

(10) Roux, B. *Curr. Opin. Struct. Biol.* **2002**, *12*, 182−189.

(11) Chang, S.-H.; Allen, T. W.; Kuyucak, S. *Biophys. J.* **2002**, *82*, 628−645.

(12) Burykin, A.; Schultz, C. N.; Villa, J.; Warshel, A. *Proteins* **2002**, *47*, 265−280.

(13) Garofoli, S.; Jordan, O. C. *Biophys. J.* **2003**, *84*, 2814−2830.

(14) Roux, B.; Berneche, S. *Biophys. J.* **2002**, *82*, 1681−1684.

(15) Allen, T. W.; Bastug, T.; Kuyucak, S.; Chung, S.-H. *Biophys. J.* **2003**, *84*, 2159−2168.

(16) Berneche, S.; Roux, B. *Nature* **2001**, *414*, 73−77.

(17) Bliznyuk, A. A.; Rendell, A. P.; Allen, T. W.; Chung, S.-H. *J. Phys. Chem. B* **2001**, *105*, 12674−12679.

(18) Cummins, P. L.; Titmuss, S. J.; Jayatilaka, D.; Bliznyuk, A. A.; Rendell, A. P.; Gready, J. E. *Chem. Phys. Lett.* **2002**, *352*, 245−251.

(19) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(20) Stewart, J. J. P. *J. Comput. Chem*. **1989**, *10*, 221−264.

(21) MOPAC2002: http://www.cachesoftware.com/mopac/index.shtml.

(22) Allen, T. W.; Bliznyuk, A.; Rendell, A. P.; Kuyucak, S.; Chung, S.-H. *J. Chem. Phys.* **2000**, *112*, 8191−8204.

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

(24) Wang, B.; Ford, G. P. *J. Comput. Chem.* **1993**, *14*, 1101−1111.

(25) Liu, C.; Walter, D.; Neuhauser, D.; Baer, P. *J. Am. Chem. Soc.* **2003**, *125*, 13936−13937.

(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; 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.; 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 B.03; Gaussian, Inc.: Pittsburgh, PA, 2003.

(27) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(28) Guidoni, L.; Carloni, P. *Biochim. Biophys. Acta* **2002**, *1563*, 1−6.

(29) Greatbanks, S. P.; Gready, J. E.; Limaye, A. C.; Rendell, A. P. *Proteins* **1999**, *37*, 157−165.

(30) Zuegg, J.; Bliznyuk, A. A.; Gready, J. E. *Mol. Phys.* **2003**, *101*, 2437−2450.

(31) Aqvist, J.; Luzhkov, V. *Nature* **2000**, *404*, 881−884.

(32) Zhang, Da. W.; Xiang, Y.; Zhang, J. Z. H. J. *Phys. Chem. B* **2003**, *107*, 12039−12041.

(33) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.

(34) Mohle, K.; Hofmann, H. J.; Thiel, W. *J. Comput. Chem.* **2001**, *22*, 509−520.

(35) Anikin, N. A.; Anisimov, V. M.; Bugaenko, V. L.; Bobrikov, V. V.; Andreyev, A. M. *J. Chem. Phys.* **2004**, *121*, 1266−1270.