Accurately Reproducing Ab Initio Electrostatic Potentials with Multipoles and Fragmentation

Hai-Anh Le, Adrian M. Lee, and Ryan P. A. Bettens*

Department of Chemistry, National University of Singapore, 3 Science Drive 3, Singapore 117543

Received: March 21, 2009; Revised Manuscript Received: July 16, 2009

In this work, we show that our energy based fragmentation method (Bettens, R. P. A.; Lee, A. M. *J. Phys. Chem. A* **2006**, 110, 8777) accurately reproduces the electrostatic potential for a selection of peptides, both charged and uncharged, and other molecules of biological interest at the solvent accessible surface and beyond when compared with the full ab initio or density functional theory electrostatic potential. We also consider the ability of various point charge models to reproduce the full electrostatic potential and compare the results to our fragmentation electrostatic potentials with the latter being significantly superior. We demonstrate that our fragmentation approach can be readily applied to very large systems and provide the fragmentation electrostatic potential for the neuraminidase tetramer (ca. 24 000 atom system) at the MP2/6-311(+)G(2d,p) level. We also show that by using at least distributed monopoles, dipoles, and quadrupoles at atomic sites in the fragment molecules an essentially identical electrostatic potential to that given by the fragmentation electrostatic potential at and beyond the solvent accessible surface can be obtained.

1. Introduction

Not surprisingly, linear scaling methods continue to attract a great deal of attention in computational chemistry. Such methods promise high accuracy at predicting energies and molecular properties that depend upon energy derivatives at a very modest cost compared to a full ab initio or density functional theory (DFT) calculation. Perhaps the earliest attempt at these methods was by Yang and Lee[1] with their density matrix divide and conquers approach. Other such methods include Kitaura's fragment molecular orbital approach (ref 2 and recently ref 3 and references therein), which has been implemented in the GAMESS package, and Exner and Mezey's field-adapted adjustable density matrix approach.[4−8]

Alternative linear scaling methods rely upon fragmenting a large molecule into smaller complete molecules or molecular systems, then linearly combining their electronic energies to approximate the total energy of the target molecule. An advantage of this approach is that standard ab initio or DFT techniques can be applied to the individual molecules, and only very course parallelization can affect a very significant speed-up in the CPU time. This can most simply be achieved by allocating a single processor to each individual complete molecular fragment system. This means that, provided enough processors exist, the time taken to compute the electronic energy of a large molecule will be as long as the most time-consuming fragment molecule calculation only. This type of approach might generally be termed molecular energy-based fragmentation.

The first type of molecular energy-based fragmentation was attempted by Gadre's group using their Molecular Tailoring Method.[9−13] This method was originally suggested to calculate one-electron properties such as the electrostatic potential. More recently Zhang and Zhang developed a quite different fragmentation algorithm[14] (Molecular Fractionation with Conjugated Caps), and have applied (and extended) their approach successfully to several systems.[15−28] This method was originally designed to accurately compute interaction energies between two molecular systems. Molecular total energies were first attempted using molecular energy-based fragmentation by Li's group[29−35] (from which we have borrowed the term molecular energy-based fragmentation), the Collins group,[36−39] and later by ourselves.[40−42]

Our method, which is quite similar to the Collins approach, produces the smallest molecular fragments of all the molecular energy-based fragmentation methods. The fragment molecules are determined entirely by the valence bonding taking place within the target molecule. Thus in the case of proteins, fragment molecules are constructed entirely from the primary sequence. Clearly nonbonded interactions play a curial role in protein interactions, and there are various ways to deal with the nonbonded interactions using our approach, but we have not as yet found one that is entirely satisfactory in all respects.

This being the case, we are able to presently deal with the most significant of nonbonded interactions, those interactions being that of net charged groups present within proteins (i.e., protonated and deprotonated side chains, metal cations, etc.). The influence of charges on the electronic structure of a protein is particularly long ranged, so it must be taken account of when computing the energy of any particular fragment molecule. This is readily achieved by placing point charges at the locations of atoms within charged groups in the protein around the fragment molecule being computed. (Here a charged group means either a $-COO^-$ or $-NH_3^+$. If such charged groups are present then an ab initio or DFT calculation was performed on $HCOO^-$ or NH_4^+ and a rank zero DMA then was performed to obtain the point charges to be placed at the positions of these nuclei. If a fragment molecule was terminated adjacent to a charged group then the fragment molecule was "grown" to include the charged group. This "charged growth" procedure has been implemented by us previously.[41]) The sign and value of these charges are determined from the rank zero multipoles (charges) extracted from Stone's distributed multipole analysis.[43−45]

The electrostatic potential around molecules, particularly those of biological interest, has been investigated by several of the

* To whom correspondence should be addressed. E-mail: chmbrpa@nus.edu.sg. Phone: +65 6516 2846. FAX: +65 6779 1691.

10.1021/jp9025706 CCC: $40.75
Published on Web 08/27/2009

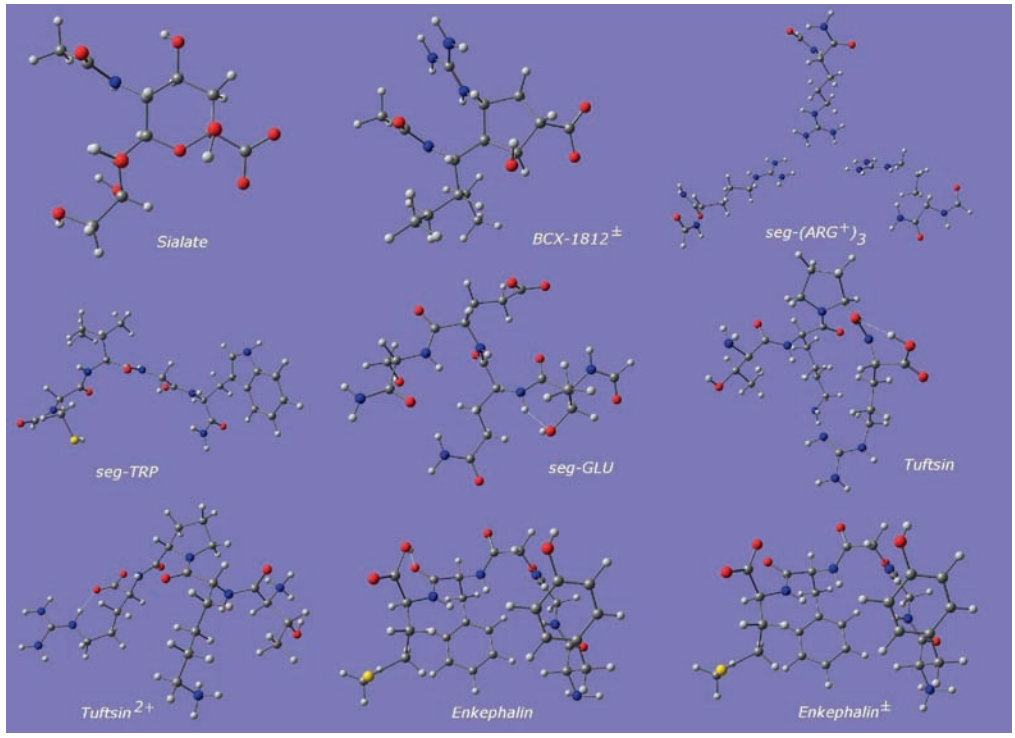

Figure 1. Systems studied in the work.

above groups working on linear scaling methods. All methods have been reported to reproduce well the electrostatic potential obtained from full calculations. However, we are yet to demonstrate that our fragmentation approach, which yields smaller fragment molecule sizes, can also accurately reproduce the electrostatic potential. This is one of the aims of the work presented here. We also compare the error associated with our computed electrostatic potential and that given by the ubiquitous point charge models. Finally, we show that our fragmentation approach can be applied to very large systems to obtain an electrostatic potential. The molecule chosen is a surface glycol protein of the influenza virus, neuraminidase, being an approximate 24 000 atom system.

2. Methods

Figure 1 shows the nine different systems studied; each of them is described in detail below. Note that all the Cartesian coordinates for each of the structures below are available in Supporting Information.

Sialate. $C_{11}H_{18}O_9N^-$. The structure was extracted from the crystal structure of sialic acid bound to neuraminidase (1MWE[46]) found in the protein data bank. The sialate selected was that bound in the active site of neuraminidase. Hydrogens were added using standard X-H bond lengths and generalized VSEPR.

BCX-1812 ± (zwitterion). $C_{15}H_{28}O_4N_4$. The structure was extracted from the crystal structure of BCX-1812 bound to neuraminidase (1L7F[47]) found in the protein data bank. Hydrogens were added using standard X−H bond lengths and generalized VSEPR.

seg-$(ARG^+)_3$. $C_{21}H_{48}O_6N_{15}^{3+}$. Arg 118, Arg 292, Arg 371 were extracted from the crystal structure, 1MWE,[46] obtained from the protein data bank. Hydrogens were added using standard X−H bond lengths and generalized VSEPR. The terminal groups were amides which were taken from the neighboring residues.

seg-TRP (Cys-Ile-Gly-Trp, residue 175−178). $C_{23}H_{32}O_5N_6S$. The structure was extracted from the crystal structure, 1L7F,[47] obtained from the protein data bank. Hydrogens were added Gaussview. The terminal groups were amides which were taken from the neighboring residues.

seg-GLU^-. $C_{18}H_{29}O_{10}N_6^-$ (Thr-Gln-Glu-Ser, residue 225−228). The structure was extracted from the crystal structure, 1MWE,[46] obtained from the protein data bank. Hydrogens were added using Gaussview. The terminal groups were amides which were taken from the neighboring residues.

Tuftsin. $C_{21}H_{40}O_6N_8$. The tetrapeptide THR-LYS-PRO-ARG. The structure was obtained from a geometry optimization at the HF/6-31G(d,p) level beginning with a structure approximating that provided in Figure 10 of ref 48.

$Tuftsin^{2+}$. $C_{21}H_{41}O_6N_8^{2+}$. As above, except the guanidinium group in ARG and the amine group in LYS were protonated. Furthermore the terminal groups were protonated and deprotonated.

Enkephalin. $C_{27}H_{35}O_7N_5S$. The pentapeptide TYR-GLY-GLY-PHE-MET. The structure was taken from the crystal structure, 1PLX,[49] obtained from the protein data bank, with hydrogens already included. No optimization was done.

Enkephalin ± (zwitterion). $C_{27}H_{35}O_7N_5S$. As above, but the terminal acid and amine groups were deprotonated and protonated respectively.

The electrostatic potentials for all the systems studied in this work were computed from the electronic wave functions obtained at various levels of theory using the *Gaussian 03* suite of programs.[50] The electrostatic potential was evaluated on a grid with a spacing of 0.2 Å out to a distance of approximately 7.5 Å from the periphery atoms on each side of the molecules studied. The various levels of theory used are summarized in Table I.

We investigated the ability to accurately reproduce the above ab initio/DFT electrostatic potential of a molecule at the specified level of theory via several different means. First, the ability of Stone's distributed multipole analysis (DMA)[44] for

TABLE I: Levels of Theory Considered for Each of the Systems Studied in This Work

system	level of theory	applied to calculation type[a]
sialate	B3LYP/6-311+G(2d,p)	all
BCX-1812 $^{\pm}$	B3LYP/6-311+G(2d,p)	all
seg-$(ARG^{+})_3$	B3LYP/6-311G(2d,p)	fragmentation
seg-TRP	HF/6-311G(d,p)	DMA, Mulliken, NPA
	B3LYP/6-311G(2p,d)	fragmentation
seg-GLU^{-}	HF/6-311+G(d,p)	all
Tuftsin	HF/6-31G(d,p)	DMA, Mulliken, NPA
	B3LYP/6-311G(2p,d)	fragmentation
$Tuftsin^{2+}$	HF/6-31G(d,p)	DMA, Mulliken, NPA
	HF/6-311+G(2p,d)	fragmentation
Enkephalin	HF/6-31G(d,p)	DMA, Mulliken, NPA
	HF/6-311G(2p,d)	fragmentation
Enkephalin$^{\pm}$	HF/6-31G(d,p)	DMA, Mulliken, NPA
	B3LYP/6-311+G(2p,d)	fragmentation

[a] See text for details.

TABLE II: Molecular System with the Level of Fragmentation and the Corresponding Total Number of Fragment Molecules, Average Number of Atoms Per Fragment Molecule and the Largest Fragment Molecule

system	level of fragmentation	total no. of fragments	average fragment size	largest fragment size (no. of heavy atoms)
sialate	2	9	13.2	24 (13)
	3	7	20.7	28 (16)
BCX-1812 $^{\pm}$	2	11	14.1	29 (14)
	3	7	24.4	36 (18)
seg-TRP	2	21	11.8	17 (9)
	3	18	16.9	24 (13)
	4	15	22.5	30 (17)
seg-GLU^{-}	2	19	11.6	21 (12)
	3	16	17.6	24 (13)
	4	13	24.1	32 (18)
Tuftsin	2	29	11.3	20 (10)
	3	26	16.1	26 (12)
	4	23	21.2	32 (14)
$Tuftsin^{2+}$	2	15	18.2	30 (11)
	3	13	22.8	38 (20)
Enkephalin	2	31	10.7	16 (8)
	3	28	14.9	20 (10)
	4	25	19.4	26 (14)
Enkephalin$^{\pm}$	2	15	13.8	31 (18)
	3	13	18.2	34 (19)
	4	13	23.3	38 (22)

reproducing the electrostatic potential was investigated. In this work, the applied DMA involved placing electrostatic multipoles on all of the nuclei within a molecule. Up to rank four multipoles (hexadecapoles) were considered. The multipoles on individual nuclei were computed from the electron density surrounding them with the density being obtained from the electronic wave function. For further detains, see ref 44. Second, we considered the distributed monopoles from the Mulliken and the Natural Population Analysis (NPA) as well as the CHARMM charges,[51] the latter only for the Enkephalin$^{\pm}$ and $Tuftsin^{2+}$ systems. These first two sets of comparisons were made without the fragmentation approximation. Next, we considered the accuracy of our fragmentation approximation[40] by linearly combining the electrostatic potentials of the fragment molecules using the fragmentation coefficients. As the quality of the fragmentation method depends upon size of the fragment molecules we have provided in Table II the number of fragment molecules and fragment size as a function of fragmentation level applied. Generally the higher the fragmentation level, the more accurate

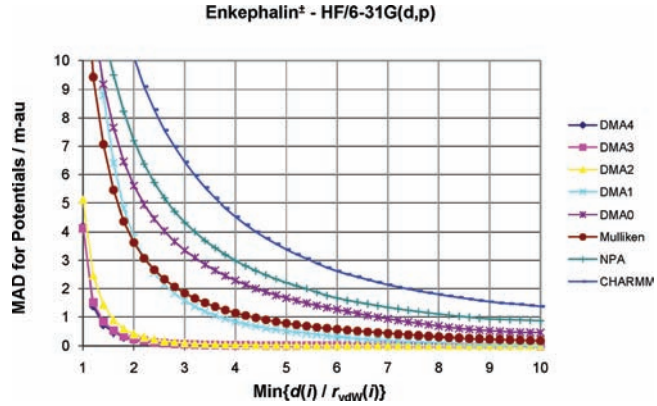

Figure 2. The mean absolute deviation (MAD/m-au) between the HF electrostatic potential and the following approximations to it: DMA*l* uses distributed multipoles up to and including rank *l*. Mulliken, NPA, and CHARMM use distributed monopoles only. The former two were determined from the so-named population analyses and the latter from ref 51.

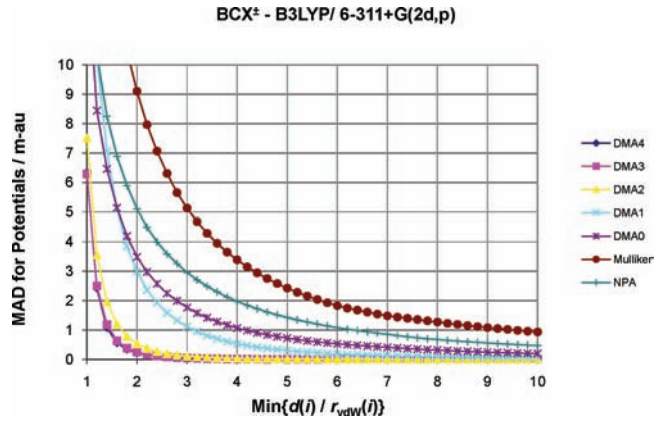

Figure 3. The mean absolute deviation (MAD/m-au) between the HF electrostatic potential and the following approximations to it: DMA*l* uses distributed multipoles up to and including rank *l*. Mulliken and NPA distributed monopoles were determined from the so-named population analyses.

the results because molecular fragments are larger. Finally, we applied the DMA to these fragment molecules, generated the associated electrostatic potentials, and again linearly combined them using the fragmentation coefficients. The levels of theory used for each of these calculation types are also listed in Table I.

3. Results and Discussion

A large amount of volumetric data was generated in determining the electrostatic potential around molecules on the above-described grids. This data is represented in Figures 2−4 by assigning any particular point on the grid to a bin. Each bin was determined by taking the minimum value from the set of ratios of the distances between the point and each atom in the molecule, $d(i)$, to the corresponding van der Waals radii, $r_{vdW}(i)$. (van der Waals radii used: $r_{vdW}(H) = 1.20$ Å, $r_{vdW}(C) = 1.70$ Å, $r_{vdW}(N) = 1.55$ Å, $r_{vdW}(O) = 1.52$ Å, $r_{vdW}(S) = 1.80$ Å.) The ab initio/DFT electrostatic potentials at each of the points in a given bin can be compared to approximations to it and summarized as a mean absolute deviation (MAD) in milliatomic units (m-au). (One m-au of electrostatic potential corresponds to ~27.2 mV or ~2.63 kJ mol^{-1} of interaction energy with a +1 electronic charge.) Figures 2−6 show these MAD versus $Min[d(i)/r_{vdW}(i)]$ for the three molecules Enkephalin$^{\pm}$, BCX$^{\pm}$,

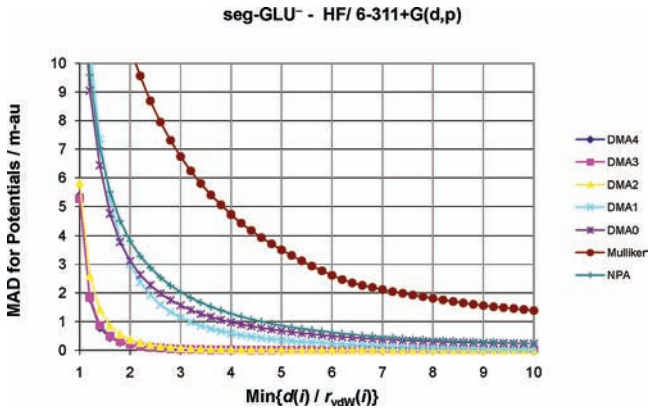

Figure 4. The mean absolute deviation (MAD/m-au) between the HF electrostatic potential and the following approximations to it: DMA*l* uses distributed multipoles up to and including rank *l*. Mulliken and NPA distributed monopoles were determined from the so-named population analyses.

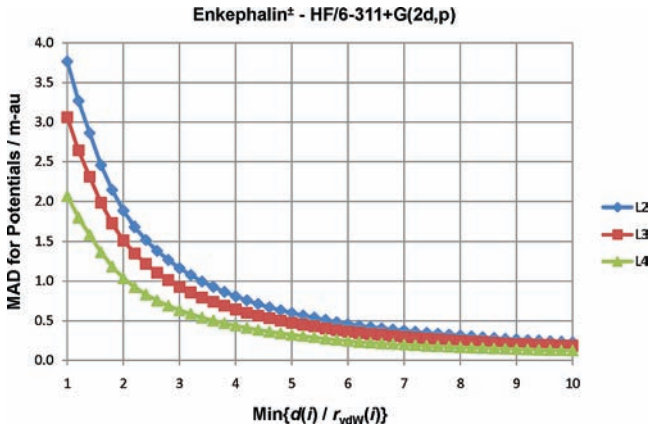

Figure 5. The mean absolute deviation (MAD/m-au) between the HF electrostatic potential and the fragmentation approximation to it. "L*n*" refers to the three different levels of fragmentation applied.

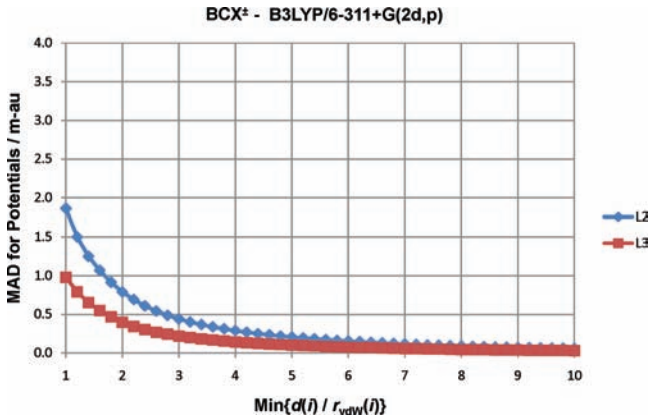

Figure 6. The mean absolute deviation (MAD/m-au) between the B3LYP electrostatic potential and the fragmentation approximation to it. "L*n*" refers to the two different levels of fragmentation applied.

and seg-GLU$^-$. The results for all the molecules studied in this work (provided in the Supporting Information) show similar trends.

It is clear by examining Figures 2−4 that the accuracy of the electrostatic potentials using only distributed monopoles (including DMA0) is, at best, qualitative; they're being wildly inaccurate at close range and fairly irreproducible in the degree of error observed. This is particularly relevant if one is solving Poisson's, or the Poisson−Boltzmann equations in order to

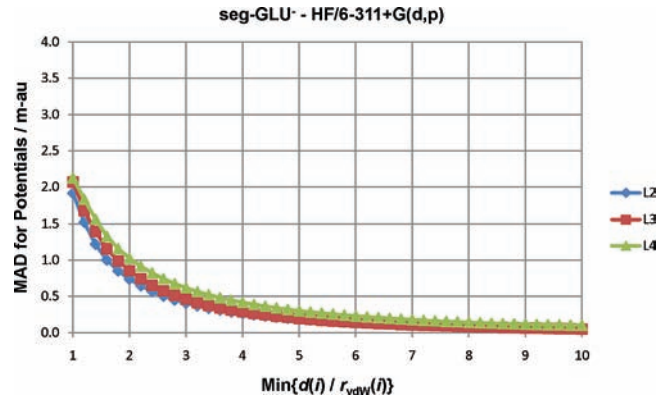

Figure 7. The mean absolute deviation (MAD/m-au) between the HF electrostatic potential and the fragmentation approximation to it. "L*n*" refers to the three different levels of fragmentation applied.

compute the electrostatic potential in an implicit solvent. Solution of these equations intimately involves the boundary conditions which can be chosen as the potential at the solvent accessible surface. In Figures 2−4, using the typical probe radius of 1.4 Å, this distance corresponds to 1.9−2.1 for Min[$d(i)/r_{vdW}(i)$], a location where the electrostatic potential is very poorly reproduced by distributed monopoles.

Similarly, DMA1 performs almost as poorly at reproducing the electrostatic potential as the distributed monopoles, although it consistently performs better than DMA0. It is not until DMA2 is reached before accurate reproduction of the potential is achieved. Indeed, at the solvent accessible surface the potential is reasonably well reproduced by distributed monopoles, dipoles, and quadrupoles with the MAD falling well below 1 m-au, and the top 10% errors at Min[$d(i)/r_{vdW}(i)$] = 2 only being ~1 m-au for all molecules. Including even higher rank multipoles only improves the accuracy.

Of note is that for overall neutral systems with nonzero central dipoles, distributed monopoles never reproduce the potential at long-range unless the distributed monopoles are altered so that they reproduce the central dipole of the molecule. This is true for DMA0 as well, but when distributed dipoles are also included (DMA1), at least the potential converges to the ab initio one at long-range. For systems with a net overall charge, all distributed monopoles treatments eventually converge to the ab initio potential, although the distance at which this occurs may be quite far (cf. Figure 6, the Mulliken analysis).

Figures 5−7 illustrate the MAD of the electrostatic potentials versus Min[$d(i)/r_{vdW}(i)$] under the fragmentation approximation for Enkephalin$^{\pm}$, BCX$^{\pm}$, and seg-GLU$^-$ respectively. Graphs for all the molecules studied in this work can be found in the Supporting Information. As expected, the results using DMA in the fragments including multipoles with ranks greater than one are essentially identical to Figures 5−7 and those provided in the Supporting Information for distances further than about 1.5 in Min[$d(i)/r_{vdW}(i)$]. Furthermore, we also note that we were able to reduce the number of sites to which the multipoles were distributed to about half by placing sites only upon heavy atoms. There was no significant loss of accuracy by performing this reduction in the number of sites.

It is clear from a comparison of Figures 5−7 with Figure 2−4 that the fragmentation approximation does not perform as well as applying DMA to the whole molecule and using multipoles with ranks greater than one at distances around Min[$d(i)/r_{vdW}(i)$] ≈ 2. However, the errors indicated in Figures 5−7 are substantially smaller than the corresponding errors resulting from distributed charges, be they either NPA, DMA0,

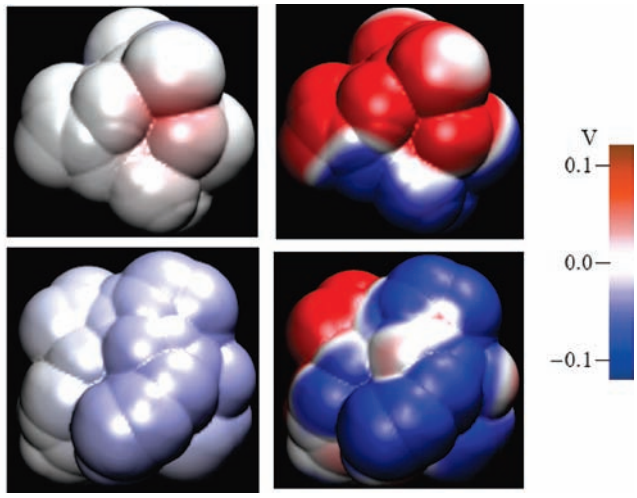

Figure 8. The error in the electrostatic potential mapped onto the solvent accessible surface for BCX$^{\pm}$ (upper panels), and Tuftsin (lower panels). The left two panels represent the error due to fragmentation approximation, whereas the right two represents the error using NPA distributed charges. For reference, the interaction energy of a 0.5*e* charge with 0.1 V potential is about 5 kJ mol^{-1}.

CHARMM, or Mulliken. Furthermore the worst error associated the fragmentation approximation at distances of Min[$d(i)/r_{vdW}(i)$] ≈ 2 is only 1 m-au, which corresponds to an interaction energy of about 1 kJ mol^{-1} for a 0.5*e* point charge. Thus the fragmentation approximation appears quantitative with respect to reproducing the electrostatic potential of the whole molecule that would be obtained from a full ab initio calculation on the same.

Figure 8 graphically illustrates the level of accuracy that can be achieved using the fragmentation approximation in comparison to NPA charges obtained from a calculation of the whole molecule. In this figure BCX and Tuftsin were selected and the error in the electrostatic potential mapped onto the solvent accessible surface (probe radius was 1.4 Å). The MADs at Min[$d(i)/r_{vdW}(i)$] = 2 are around 11, 27, 139, 115 mV for L3 fragmentation of BCX, L4 fragmentation of tuftsin, and NPA charges of BCX and tuftsin respectively.

We would like to note at this point that the accuracy of the distributed charges can be improved considerably by fitting a set of charges by minimizing the sum of the mean square deviation between the distributed-charge generated potential and the actual potential. However, this is clearly not an option if the molecule under study is a large protein or enzyme as the actual potential is not available.

To illustrate the viability of generating an electrostatic potential for an enzyme using fragmentation, we have fragmented the neuraminidase monomer at level 2 (as described previously[41]). Neuraminidase is a surface glycoprotein of the influenza virus and is actually a tetramer with C_4 symmetry containing about 24 000 atoms. The heavy atom structure was abstracted from the Protein Data Bank, 1MWE.[46] Thus using the symmetry of the system, we generated the electrostatic potential of which an isosurface is illustrated in Figure 9. The level of theory used to generate the electrostatic potential was MP2/6-311(+)G(2d,p) (MP2 density was used).

Although the electrostatic potential illustrated is fully quantum mechanical, no solvent, explicit or implicit, was used in generating it. Therefore while the potential is expected to be accurate in vacuo, it can only at best be viewed as qualitative in regard to solvated neuraminidase. Nevertheless there are some interesting features that, perhaps, warrant mention.

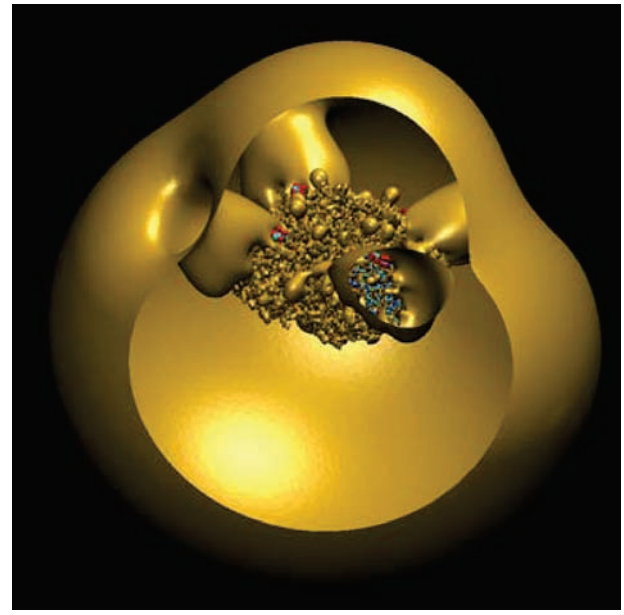

Figure 9. Electrostatic potential isosurface for the neuraminidase tetramer (about −0.5 V) computed at the MP2/6-311(+)G(2d,p) level of theory. The neuraminidase tetramer possess about 24 000 atoms. The siliates bound in the active site and secondary binding site of each monomer can just be made out.

The charge state assumed for the electrostatic potential calculation is that "expected" (using standard pK_a's) for neuraminidase at pH = 7, that is, −4, although each monomer contained a total of 86 charged groups including a native Ca^{2+} species. The species of natural action of neuraminidase is a sialoglycoconjugate found on the surface of target cells - an anion at pH = 7. It is observed that for small negative values of the potential an approximately spherical isosurface encloses the entire tetramer, that is, at very long-range neuraminidase appears as a point −4 charged species. However, it is fascinating to observe that by making the potential more negative a "tunnel" opens up in the isosurface which leads only to the active site and secondary binding sites of sialate in neuraminidase. The four tunnels of the tetramer are clearly visible in the cutaway shown in Figure 9. It is of note that this feature is not at all apparent by examination of the electrostatic potential of the monomer only. Also of note is that removal of the four native Ca^{2+} destroys this scenario with the active site no longer at all obvious in the isosurfaces. Finally, a fifth Ca^{2+} is often found in crystal structures to lie along the C_4 axis and appears to be a counterion or Ca^{2+} of crystallization. The location with the most negative potential along the C_4 axis was found to lie within 1 Å of this Ca^{2+}. Close inspection of the crystal structure revealed that this Ca^{2+} is surrounded by what appears to be a hydration shell reminiscent of von Helmholtz's double layer.[52] Also of note is that when we included this Ca^{2+} in the electrostatic potential calculation the four tunnels exhibited in Figure 9 were still present, although they were rotated to a more upward facing position.

Even though the electrostatic potential computed in this work is fully quantum mechanical, it is still in vacuo. While crude, it is considered by many that a better electrostatic potential could be obtained though solution of the Poisson−Boltzmann equation using some impalpable ad hoc distance dependent dielectric constant. Such an approach could be achieved using fragmentation by making the quantum mechanical electrostatic potential at the solvent accessible surface a boundary condition for solution of the Poisson−Boltzmann equation.

An alternative approach to obtain an electrostatic potential that incorporates the effects of the solvent, at least implicitly, would be to use one of the myriad of solvation models currently available in standard ab intio/DFT packages on the fragment molecules. Preliminary tests by us on small molecules reveal that the fragmentation approximation performs equally well at reproducing total electronic energies using a solvation model as it does in vacuo. However, for very large systems like enzymes questions arise as to the appropriate cavity to embed the fragment molecule within; the cavity should not reflect the shape of the fragment molecule itself, but rather the shape of the enzyme in which the fragment molecule is a part. For very large molecules like enzymes the difference between these two cavities is pronounced and would thus be expected to affect the results markedly.

Regardless of the issues associated with the use of continuum models, there is another much more fundamental and serious question to which its answer is expected to have a considerable impact on any computed electrostatic potential. That question being the following: What is the actual protonation state of a solvated protein anyway? Quite obviously a close contact carboxylate and protonated amine will not exist in nature for more than a few molecular vibrations in any phase of matter, yet this is the standard assumption made in many electrostatic potential calculations (including this work!). Similarly, like charged groups in close proximity to each other (e.g., seg-$(ARG^+)_3$ system which "is present" in the active site of neuraminidase) are unlikely to exist in the protonated/deprotonated state exhibited by the same groups far from each other and fully exposed to solvent. Clearly, the impact on the electrostatic potential will be substantial if one assumes, effectively, a large dipole for adjacent oppositely charged groups compared to simply H-bonded species, or three close positive charges when most of the time there is only two or fewer.

4. Conclusions

We have shown that our fragmentation method is accurate at reproducing the electrostatic potential of various systems. Additionally, our method is significantly superior to standard point charge models of the same systems. Essentially identical results at distances further than 1.5 times the van der Waals radii can be obtained if Stone's distributed multipoles are utilized. The significance of the latter is as follows. Most proteins and enzymes are made up of about 21 amino acids. Our fragmentation method fragments the target molecule based upon primary sequence only. So while the monomer of neuraminidase fragments into 2272 different molecular structures at level 2, only 46 of these represents parent structures with all the remaining structures being simple conformations of these 46 parents. Stone's recent DMA algorithm[44] produces highly transferable distributed multipoles, so in principle it is possible to produce a database of fragment molecular structures with associated multipoles (and polarizabilities to account for the necessary charge field around fragments) and perhaps a possible parameter set to account for any significant variation in the multipoles with conformation. Such a database might then be applied to any protein structure. We have also shown that our method can be readily applied to very large systems like the neuraminidase tetramer, an approximate 24 000 atom system. Further work is needed to incorporate the effects of the solvent into the computation.

Acknowledgment. We thank the National University of Singapore for the faculty research Grant (R-143-000-262-112) that supported this work, and the Centre of Computational Science and Engineering for the kind use of their computing facilities.

Supporting Information Available: Cartesian coordinates of sialate, BCX-1812$^{\pm}$, seg-$(ARG^+)_3$, seg-TRP, seg-GLU^-, Tuftsin, Tuftsin^{2+}, Enkephalin, and Enkephalin$^{\pm}$. Figures including DMA, Mulliken, and NPA for sialate, BCX-1812$^{\pm}$, seg-$(ARG^+)_3$, seg-TRP, seg-GLU^-, Tuftsin, Tuftsin^{2+}, Enkephalin, and Enkephalin$^{\pm}$. Figures including the fragmentation results for sialate, BCX-1812$^{\pm}$, seg-$(ARG^+)_3$, seg-TRP, seg-GLU^-, Tuftsin, Tuftsin^{2+}, Enkephalin, Enkephalin$^{\pm}$. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Yang, W. T.; Lee, T. S. *J. Chem. Phys.* **1995**, *103*, 5674.
(2) Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2000**, *318*, 614.
(3) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904.
(4) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2004**, *108*, 4301.
(5) Exner, T. E.; Mezey, P. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 4061.
(6) Eckard, S.; Exner, T. E. *Z. Phys. Chem.* **2006**, *220*, 927.
(7) Exner, T. E.; Mezey, P. G. *J. Comput. Chem.* **2003**, *24*, 1980.
(8) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2002**, *106*, 11791.
(9) Rahalkar, A. P.; Ganesh, V.; Gadre, S. R. *J. Chem. Phys.* **2008**, *129*, 234101.
(10) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. *J. Chem. Phys.* **2006**, *125*, 104109.
(11) Gadre, S. R.; Shirsat, R. N.; Limaye, A. C. *J. Phys. Chem.* **1994**, *98*, 9165.
(12) Babu, K.; Gadre, S. R. *J. Comput. Chem.* **2003**, *24*, 484.
(13) Babu, K.; Ganesh, V.; Gadre, S. R.; Ghermani, N. E. *Theor. Chem. Acc.* **2004**, *111*, 255.
(14) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599.
(15) Mei, Y.; Ji, C. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 7.
(16) Chen, X. H.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 7.
(17) Chen, X. H.; Zhang, Y. K.; Zhang, J. Z. H. *J. Chem. Phys.* **2005**, *122*, 5.
(18) Zhang, D. W.; Zhang, J. Z. H. *Int. J. Quantum Chem.* **2005**, *103*, 246.
(19) He, X.; Zhang, J. Z. H. *J. Chem. Phys.* **2005**, *122*, 4.
(20) Mei, Y.; Zhang, D. W.; Zhang, J. Z. H. *J. Phys. Chem. A* **2005**, *109*, 2.
(21) Chen, X. H.; Zhang, J. Z. H. *J. Theory Comput. Chem.* **2004**, *3*, 277.
(22) Gao, A. M.; Zhang, D. W.; Zhang, J. Z. H.; Zhang, Y. K. *Chem. Phys. Lett.* **2004**, *394*, 293.
(23) Xiang, Y.; Zhang, D. W.; Zhang, J. Z. H. *J. Comput. Chem.* **2004**, *25*, 1431.
(24) Chen, X. H.; Zhang, J. Z. H. *J. Chem. Phys.* **2004**, *120*, 11386.
(25) Zhang, D. W.; Xiang, Y.; Gao, A. M.; Zhang, J. Z. H. *J. Chem. Phys.* **2004**, *120*, 1145.
(26) Chen, X. H.; Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2004**, *120*, 839.
(27) Zhang, D. W.; Xiang, Y.; Zhang, J. Z. H. *J. Phys. Chem. B* **2003**, *107*, 12039.
(28) Zhang, D. W.; Chen, X. H.; Zhang, J. Z. H. *J. Comput. Chem.* **2003**, *24*, 1846.
(29) Li, S. H.; Li, W.; Fang, T. *J. Am. Chem. Soc.* **2005**, *127*, 7215.
(30) Dong, H.; Hua, S. G.; Li, S. H. *J. Phys. Chem. A* **2009**, *113*, 1335.
(31) Li, W.; Dong, H.; Li, S. H. In *Relative Energies of Proteins and Water Clusters Predicted with the Generalized Energy-Based Fragmentation Approach*, Proceedings of the 12th European Workshop on Quantum Systems in Chemistry and Physics, August 30−September 5, 2007; Univ London, *Royl. Holloway Coll.*, London, England. Source: *Frontiers in Quantum Systems in Chemistry and Physics Book Series: Progress in Theoretical Chemistry and Physics.* Volume: 18, pps: 289−299, published: 2008.
(32) Hua, W. J.; Fang, T.; Li, W.; Yu, J. G.; Li, S. H. *J. Phys. Chem. A* **2008**, *112*, 10864.
(33) Li, H.; Li, W.; Li, S. H.; Ma, J. *J. Phys. Chem. B* **2008**, *112*, 7061.
(34) Li, W.; Li, S. H.; Jiang, Y. S. *J. Phys. Chem. A* **2007**, *111*, 2193.
(35) Li, W.; Fang, T.; Li, S. H. *J. Chem. Phys.* **2006**, *124*, 6.
(36) Deev, V.; Collins, M. A. *J. Chem. Phys.* **2005**, *122*, 154102.
(37) Collins, M. A.; Deev, V. A. *J. Chem. Phys.* **2006**, *125*, 104104.
(38) Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 10.
(39) Netzloff, H. M.; Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 13.
(40) Bettens, R. P. A.; Lee, A. M. *J. Phys. Chem. A* **2006**, *110*, 8777.
(41) Bettens, R. P. A.; Lee, A. M. *Chem. Phys. Lett.* **2007**, *449*, 341.

(42) Lee, A. M.; Bettens, R. P. A. *J. Phys. Chem. A* **2007**, *111*, 5111.
(43) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233.
(44) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128.
(45) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047.
(46) Varghese, J. N.; Colman, P. M.; vanDonkelaar, A.; Blick, T. J.; Sahasrabudhe, A.; McKimmBreschkin, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11808.
(47) Smith, B. J.; McKimm-Breshkin, J. L.; McDonald, M.; Fernley, R. T.; Varghese, J. N.; Colman, P. M. *J. Med. Chem.* **2002**, *45*, 2207.
(48) Rafat, M.; Popelier, P. L. A. *J. Comput. Chem.* **2007**, *28*, 832.
(49) Marcotte, I.; Separovic, F.; Auger, M.; Gagne, S. M. *Biophys. J.* **2004**, *86*, 1587.
(50) Frisch, M. J; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; J. A. Montgomery, J.; 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*, rev. D.01; Gaussian, Inc.: Wallingford, CT, 2004..
(51) Charifson, P. S.; Hiskey, R. G.; Pedersen, L. G. *J. Comput. Chem.* **1990**, *11*, 1181.
(52) von Helmholtz, H. *Ann. Phys. Chem.* **1853**, *89*, 211.

JP9025706