JCTC Journal of Chemical Theory and Computation

Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional

Carlos Campañá, Bastien Mussard, and Tom K. Woo*

Centre for Catalysis Research and Innovation and Department of Chemistry, University of Ottawa, Ottawa, Canada, K1N 6N5

Received July 7, 2009

Abstract: A method to generate electrostatic potential (ESP) derived atomic charges in crystalline solids from periodic quantum mechanical calculations, termed the REPEAT method, is presented. Conventional ESP fitting procedures developed for molecular systems, in general, will not work for crystalline systems because the electrostatic potential in periodic systems is ill-defined up to a constant offset at each spatial position. In this work the problem is circumvented by introducing a new error functional which acts on the relative differences of the potential and not on its absolute values, as it is currently done with molecular ESP charge derivation methods. We formally demonstrate that the new functional reduces to the conventional error functional used in molecular ESP approaches when the simulation box of the periodic calculation becomes infinitely large. Several tests are presented to validate the new technique. For the periodic calculation of isolated molecules, the REPEAT charges are found to be in good agreement with those determined with established molecular ESP charge derivation methods. For siliceous sodalite, it is demonstrated that conventional molecular ESP approaches generate 'unphysical' charges, whereas the REPEAT method produces charges that are both chemically intuitive and consistent between different periodic electronic structure packages. The new approach is employed to generate partial atomic charges of various microporous materials and compared to both experimentally derived and molecular fragment ESP charges. This method can be used to generate partial atomic charges to be used in simulations of microporous and nanoporous materials, such as zeolites and metal organic framework materials.

I. Introduction

Microporous materials are a fascinating class of condensed matter systems with a wide range of important applications. They are commonly employed as catalysts in the petrochemical industry, as adsorbents to trap impurities, as physical reservoirs for gas storage, as molecular sieves for size and shape selective separations, and as mediums for ion exchange processes.[1] Perhaps the most well-known microporous materials are zeolites, which are crystalline solids with well-defined structures and pores that range in size from $\sim$1−20 Å.[1] Recently, a new class of microporous materials known as metal organic frameworks (MOFs)[2] have emerged that are composed of metal ions and bridging organic ligands. Compared to zeolites, MOFs have a much wider variety of structural and chemical motifs that promise a higher degree of tunability. As such, MOFs have attracted significant attention for their potential applications in hydrogen storage and CO_2 sequestration.[3−9]

Atomistic modeling of microporous materials, particulary zeolites, have contributed enormously to increase our understanding on their properties and functionality.[10,11] Most studies of the full periodic systems have employed empirical interatomic potentials (force fields), although periodic density functional theory (DFT) studies have recently emerged.[11] Molecular dynamics and Monte Carlo simulations using empirical potentials have been successfully applied to investigate the dynamics of guest molecules as well as their adsorption characteristics within microporous materials.[10−13] In some cases, atomistic simulations have led to the design of improved compounds with enhanced properties and

* Corresponding author. E-mail: twoo@uottawa.ca.

10.1021/ct9003405 CCC: $40.75
Published on Web 08/27/2009

performance.[14,15] Although several generally transferable force fields have been developed for organic molecules, the treatment of electrostatic interactions still remains a challenge when employing empirical potentials. Most force fields use fixed partial atomic charges to treat the electrostatic interactions, although more sophisticated methods have been developed but are not in wide use. Since there are no 'true' atomic charges within polynuclear systems, an assortment of charge derivation methods have been created for various quantitative and descriptive purposes. The so-called electrostatic potential (ESP) derived charges are most commonly used for atomistic simulation of molecular systems.[16−19] In order to compute ESP charges, a quantum chemical calculation is performed on a molecule, and partial atomic charges are fit to reproduce the quantum chemical electrostatic potential on a fine grid surrounding the molecule. The grid is chosen to lie outside of the van der Waals radius (or similar) of each atom in the molecule. The charges are 'designed' to reproduce the electrostatic potential in the region where it is most important when modeling intermolecular interactions with two-body, additive Coulomb potentials. In situations where there are deeply embedded (buried) atoms in the molecule, the ESP charges for these buried atoms can fluctuate widely, sometimes resulting in nonchemically intuitive values. To deal with this problem, Bayly and co-workers introduced the RESP method,[18] where a penalty function is introduced to the fitting procedure to inhibit the 'unphysical' charges from arising. ESP charges have been shown to provide superior results compared to those using other schemes, such as charges derived from a Mulliken population analysis, and therefore, ESP charges are the norm for molecular simulation when employing point charge electrostatic models.[16,18]

For the atomistic simulation of periodic systems, ESP charges are not as universally applied as they are for molecular systems. In highly packed solids, where there is little volume outside of the van der Waals radii of the atoms to define valid fitting points, ESP charges may not be appropriate. However, even for the simulation of microporous materials where there are large pore volumes, Mulliken charges derived from periodic DFT calculations are still in wide usage despite the strong basis set dependence of the scheme.[13,20] If ESP charges are used for periodic simulations, they are derived by extracting a fragment of the periodic lattice and performing a molecular ESP calculation.[12,21−23] The assumption with this procedure is that the electrostatic potential surrounding the isolated fragment is going to be similar to that of the periodic structure. Considering the care taken to treat long-range electrostatic interactions when using periodic boundary conditions in molecular simulations, this assumption may not be generally valid. Additionally, special considerations also have to be made to 'cap' the electronic system to satisfy unfilled valences resulting from the extraction or account for the net charge of the retrieved fragment. Finally, the properties of a periodic wave function including the electrostatic potential are averages over the Brillouin zone. Molecular fragment calculations cannot capture the effect of proper K-point sampling, which is important for some periodic systems.

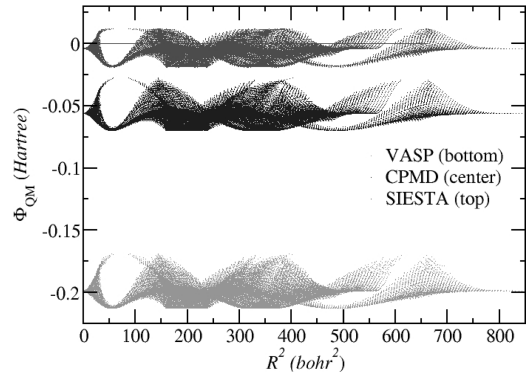

Figure 1. Electrostatic potential plotted as a function of the square of the distance from the cell origin for sodalite generated from three periodic DFT packages: VASP (bottom), CPMD (center), and SIESTA (top). The tabulation is performed on a cubic grid, and only points located outside the van der Waals radii centered on each atom are shown.

Although periodic 'first principles' (i.e., DFT) calculations of materials with large unit cells containing hundreds of atoms are now routinely performed, to the best of our knowledge, a method to derive charges from the electrostatic potential of periodic first principles calculations has not been reported.[24,25] One of the reasons for the dearth of development of periodic ESP fitting techniques, as compared to molecular systems, may be due to the fact that the absolute energy of an atom is not an intrinsic bulk property, and therefore, the average electrostatic potential is ill-defined.[26] In other words, the reference state of the electrostatic potential in a periodic system is arbitrary. This is numerically demonstrated in Figure 1, where the electrostatic potential for siliceous sodalite (a zeolite) is plotted as a function of the square distance to the cell origin on a mesh resulting from three different periodic DFT packages − VASP, CPMD, and SIESTA.[27] Qualitatively, the general structure of the electrostatic potentials show the same features, except that the manifold derived from each periodic DFT package is shifted by a different but constant offset amount. Since the average of the electrostatic potential in an infinite system is ill-defined, there is no 'correct' offset value, and the straightforward application of the ESP charge fitting procedures developed for molecular systems will result in charges that are dependent on this arbitrary offset value.

In this work, we introduce a modified ESP error functional that can be used in periodic systems as well as in molecular systems. Our functional allows for a robust computation of the ESP charges while overcoming the fundamental problem of the electrostatic potential having an ill-defined reference state. We also introduce a new RESP-like penalty function to prevent large fluctuations in the fitted charges of buried atoms that is based on the expansion of the energy of an atom as a function of charge. In the tradition of ESP charge methods, we also suggest an acronym for the charges derived from the method. We refer to these charges as REPEAT (Repeating Electrostatic Potential Extracted ATomic) charges. The remainder of the paper is as follows: The formalism of the method is introduced in detail with the appropriate equations in Section II. Section III describes the choice of parameters used in our implementation of the REPEAT

method as well as details on the periodic DFT calculations that were performed to generate the electrostatic potentials. Section IV contains the tests and the applications used to validate the REPEAT method, and Section V concludes the paper.

II. Methodology

For molecular systems, the electrostatic potential at a point $\vec{r}$ that results from a set of N_q point charges, $\{q_j\}$, is given in eq 1 (in atomic units), where $\vec{r}_j$ represents the position of each of the point charges:

$$\phi_q(\vec{r}) = \sum_j^{N_q} \frac{q_j}{|\vec{r} - \vec{r}_j|} \tag{1}$$

In generating ESP charges for molecular systems, a quantum chemical calculation is first performed to generate a reference electrostatic potential, $\phi_{QM}(\vec{r})$. The set of point charges $\{q_j\}$ is then adjusted to minimize the differences between the 'quantum mechanical' electrostatic potential, $\phi_{QM}(\vec{r})$, and the electrostatic potential due to the atomic charges defined by eq 1, for grid points surrounding the atoms in the molecule. This is typically performed in a least-squares manner where the following function is minimized:

$$F(\{q_j\}) = \sum_{grid} (\phi_{QM}(\vec{r}_{grid}) - \phi_q(\vec{r}_{grid}))^2 + \lambda(\sum_j q_j - q_{tot}) \tag{2}$$

The second term in eq 2 represents a Lagrange multiplier that ensures that the sum of atomic charges equals the total charge of the system, q_{tot}. The grid points used for the fitting procedure are chosen to lie outside of the van der Waals (VDW) radii (or similar) of the atoms in the molecule. The actual algorithm for choosing the grid points varies between different methods, and several ESP charge fitting techniques are in wide usage, such as the CHELPG[17] and the RESP[18] approaches.

When generating ESP derived atomic charges for periodic systems, there are additional complications that need to be addressed as compared to molecular ESP charges. First, the electrostatic potential due to the point charges has to account for the infinite periodic images of the atomic charges. The Ewald summation technique is a well-known mathematical procedure to deal with the long-range electrostatic interactions in periodic systems. Within the Ewald procedure, the electrostatic potential is factored in two terms, real and reciprocal space series. The real space series accounts for the effects of the short-range interactions. Conversely, the reciprocal space term includes the long-range effects arising from the presence of the infinite images. For a formal derivation of the mathematical expressions representing each one of the series, we refer the reader to the original and related works.[28,29] Summarizing, the exact expression for the potential according to the Ewald formulation is given in eq 3 where the erfc stands for the complementary error function.

$$\phi_q(\vec{r}) = \sum_{j,\vec{T}} q_j \frac{\mathrm{erfc}(\alpha|\vec{r} - \vec{r}_{j,\vec{T}}|)}{|\vec{r} - \vec{r}_{j,\vec{T}}|} + \frac{4\pi}{V} \sum_{j,\vec{k}} q_j e^{i\vec{k}(\vec{r}-\vec{r}_j)} \frac{e^{-\frac{k^2}{4\alpha^2}}}{k^2} \tag{3}$$

The vectors $\vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$, appearing in the first double summation, map the real space lattice positions. Conversely, the vectors $\vec{k} = m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3$, included within the second double summation, map the reciprocal space lattice. The variable α represents the width of the screening Gaussian charge distribution added to each atomic center. Equation 3 can be considered the periodic analogue of eq 1. With infinite terms, eq 3 is mathematically exact and independent of the value chosen for the parameter α. However, the value of α as well as the number of terms in the real and reciprocal space series determine the speed of convergence and accuracy of the numerical computation of the electrostatic potential. Techniques to select optimum values for these parameters have been discussed elsewhere.[30] If the charges are to be used in a molecular dynamics simulation, then it can be argued that these Ewald summation parameters used to derive the ESP charges should ideally be the same as those used in the simulations.

The second complication when deriving periodic ESP charges relates to the ill-defined nature of the reference state of the electrostatic potential, as depicted in Figure 1. Even if eq 3 is used for ϕ_q, minimization of the conventional error function (eq 2) developed for molecular systems will not be of any value when attempting to compute charges in periodic systems. For example, it is clear that the different electrostatic potentials shown in Figure 1 will yield drastically different fitted charges. (There are reports of conventional ESP charge fitting procedures yielding unsatisfactory results for periodic systems.)[31,32] To overcome this fundamental problem, we introduce a modified error functional to perform the least-squares fit. Our modified error functional can be expressed as the sum of two independent terms:

$$F(\{q_j, \delta_\phi\}) = \sum_{grid} (\phi_{QM}(\vec{r}_{grid}) - (\phi_q(\vec{r}_{grid}) + \delta_\phi))^2 + \lambda(\sum_j q_j - q_{tot}) \tag{4}$$

where the new parameter δ_ϕ is given by

$$\delta_\phi = \frac{1}{N_{grid}} \sum_{grid} (\phi_{QM}(\vec{r}_{grid}) - \phi_q(\vec{r}_{grid})) \tag{5}$$

As in molecular approaches, the second term in eq 4 represents a Lagrange multiplier that ensures that the total charge of the system remains equal to q_{tot}, which is zero in the case of periodic systems. Examination of the definition of δ_ϕ in eq 5 reveals that the ESP fit is performed such that only the relative differences with respect to the average value of the potential on the grid points are considered. This automatically guarantees that the quality of the fit is not affected by a constant offset in the estimation of the potential values. It is important to note that the expression for δ_ϕ can be obtained in a simple and straightforward manner by assuming δ_ϕ in eq 4 is another independent variable in the error functional $F(\{q_j, \delta_\phi\})$. The definition of δ_ϕ in eq 5 can

then be derived by performing the minimization of F not only with respect to the charges but also with respect to δ_ϕ.

Explicit consideration of δ_ϕ within the error functional form not only removes the artifact of an ill-defined reference state but implicitly introduces corrections to the dipole moment and the other higher-order multipoles, if our technique is to be employed to fit charges in molecular compounds. More specifically, it is easy to show that for a neutral molecular system if one expands δ_ϕ in the continuum representation up to the first-order (dipole) corrections one obtains

$$\delta_\phi = \frac{1}{N_{\text{grid}}}\sum_{\text{grid}}(\phi_{\text{QM}}(\vec{r}_{\text{grid}}) - \phi_q(\vec{r}_{\text{grid}})) \approx \Delta\phi_{q-\text{QM}} + \frac{1}{V}\int \frac{(\vec{p}_{\text{QM}} - \vec{p}_q)\hat{r}}{|\vec{r}|^2}d^3r \quad (6)$$

where $\vec{p}$ is the electric dipole moment and $\Delta\phi_{q-\text{QM}}$ is just the difference between the zero levels of the quantum and the classical charge distribution, which is expected to be zero in molecular systems. The second term reveals that errors in the dipole moment can be lowered using our fitting technique. The possibility of including implicit corrections to multipole moments allows for capturing potential polarization effects in the cluster calculations of charges within highly polar molecular systems.

In systems in which there are buried atoms, such that there are few valid fitting grid points close to these atoms, the ESP charge fitting approach can yield charges that are considered 'unphysical' or chemically nonintuitive. To avoid wide fluctuations in the fitted charges for buried atoms, penalty functions are added to the error functional, such as those used in the RESP method.[18] In order to generalize our methodology to systems in which buried atoms could be present, we propose adding a second set of Lagrange multipliers, w_j, such that the REPEAT error functional becomes:

$$F(\{q_j,\delta_\phi\}) = \sum_{\text{grid}}(\phi_{\text{QM}}(\vec{r}_{\text{grid}}) - (\phi_q(\vec{r}_{\text{grid}}) + \delta_\phi))^2 + \lambda(\sum_j q_j - q_{\text{tot}}) + \sum_j w_j\left(E_j^0 + \chi_j q_j + \frac{1}{2}J_j^{00}q_j^2\right) \quad (7)$$

As opposed to the original RESP method and derivative works,[18] we have made no ad hoc assumptions regarding the shape of the penalty term. We use a physically motivated penalty multiplier that is an expansion of the energy of an atom (idem to the potential in atomic units) as a function of the atomic charges up to the second order.[33,34] The parameters χ_j and J_j^{00} are the electronegativity and the self-Coulomb interaction, respectively, of the individual chemical elements. The nature of the self-Coulomb interaction term $J_j^{00}$and how to determine it is discussed in more detail by Rappe and Goddard.[34] The free parameter included in the model is w_j, which can be considered a weighting factor used to adjust the strength of the restraints. A single weighting factor w could be used for all atoms, or as expressed in eq 7, the weighting factors can be individually adjusted to selectively turn on penalty corrections only for atoms identified as buried prior to performing the charge calculations. Aside from being a physically motivated restraining method, the method also has the advantage that the derivative of the error function with respect to the charges is linear (vide infra), and therefore, there is no need to solve the equations iteratively as is required with the hyperbolic constraints of RESP.

In the following, we provide the full derivation of the linear system of equations to be solved in order to calculate atomic charges from the spatial tabulation of the periodic electrostatic potential within the REPEAT approach. Minimizing the error functional, F, defined in eq 7 with respect to the independent variables $\{q_j\}$ and δ_ϕ yields the set of equations:

$$\begin{aligned}\frac{\partial F}{\partial q_{j=1,\ldots,N_q}} &= -2\sum_{\text{grid}}(\phi_{\text{QM}}(\vec{r}_{\text{grid}}) - \delta_\phi - \phi_q(\vec{r}_{\text{grid}}))\times \left(\frac{\partial\phi_q(\vec{r}_{\text{grid}})}{\partial q_j} + \frac{\partial\delta_\phi}{\partial q_j}\right) + \lambda + w_j(\chi_j + J_j^{00}q_j) = 0,\\ \frac{\partial F}{\partial\delta_\phi} &= -2\sum_{\text{grid}}(\phi_{\text{QM}}(\vec{r}_{\text{grid}}) - \delta_\phi - \phi_q(\vec{r}_{\text{grid}})) = 0\end{aligned} \quad (8)$$

The last equation above defines δ_ϕ as presented in eq 5. Using the definition of the derivative of the potential with respect to the charges:

$$\frac{\partial\phi_q(\vec{r}_{\text{grid}})}{\partial q_j} = \sum_{\vec{T}}\frac{\text{erfc}(\alpha|\vec{r}_{\text{grid}} - \vec{r}_{j,\vec{T}}|)}{|\vec{r}_{\text{grid}} - \vec{r}_{j,\vec{T}}|} + \frac{4\pi}{V}\sum_{\vec{k}} e^{i\vec{k}(\vec{r}_{\text{grid}}-\vec{r}_j)}\frac{e^{-\frac{k^2}{4\alpha^2}}}{k^2} = \phi_j'(\vec{r}_{\text{grid}}) \quad (9)$$

combined with the first group of eq 8 and after some lengthy algebra one arrives at a matrix problem with $N_q + 1$ unknowns of the type:

$$\mathbf{A}\tilde{\mathbf{Q}} = \mathbf{B} \quad (10)$$

with matrix coefficients $\mathbf{A}_{jm}$ defined as

$$\mathbf{A}_{jm,(j,m)<N_q+1} = \sum_{\text{grid}}\left\{\left(\phi_j'(\vec{r}_{\text{grid}}) - \frac{1}{N_{\text{grid}}}\sum_{\text{grid}*}\phi_j'(\vec{r}_{\text{grid}*})\right)\times (\phi_m'(\vec{r}_{\text{grid}}) - \frac{1}{N_{\text{grid}}}\sum_{\text{grid}*}\phi_m'(\vec{r}_{\text{grid}*}))\right\} + \frac{w_j}{2}J_j^{00}\delta_{jm} \quad (11)$$

and

$$\mathbf{A}_{jm,j<N_q+1,m=N_q+1} = \mathbf{A}_{jm,j=N_q+1,m<N_q+1} = \mathbf{A}_{jm,j=m=N_q+1} = 1 - \delta_{jm} \quad (12)$$

The vector $\mathbf{B}$ on the right-hand side of eq 10 has elements

$$\mathbf{B}_{m,m<N_q+1} = \sum_{\text{grid}}\left\{(\phi_{\text{QM}}(\vec{r}_{\text{grid}}) - \frac{1}{N_{\text{grid}}}\sum_{\text{grid}*}\phi_{\text{QM}}(\vec{r}_{\text{grid}*}))(\phi_m'(\vec{r}_{\text{grid}}) - \frac{1}{N_{\text{grid}}}\sum_{\text{grid}*}\phi_m'(\vec{r}_{\text{grid}*}))\right\} - w_m\frac{\chi_m}{2} \quad (13)$$

and

$$\mathbf{B}_{m,m=N_q+1} = q_{\text{tot}} \quad (14)$$

The elements of the unknown vector $\tilde{\mathbf{Q}}$ are given by:

$$\tilde{\mathbf{Q}}_{m,m<N_q+1} = q_m \tag{15}$$

and

$$\tilde{\mathbf{Q}}_{m,m=N_q+1} = \frac{\lambda}{2} \tag{16}$$

In the case where the crystallographic symmetry of the system demands the equivalency of some of the atomic sites and therefore charges, a minor modification needs to be added to the above formulation. The columns and rows corresponding to identical sites in the **A** matrix (rows in the case of the right-hand side vector **B**) must be combined into a single unit by adding their elements.

III. Computational Details

Due to the numerical nature of this work, in this section we provide a discussion of the parameters chosen in the different tests that were performed in order to validate our methodology. We also provide the details of the DFT calculations used to generate the electrostatic potentials of the different systems considered as well as the details of other quantum chemical calculations used for comparison purposes. Some information on our implementation and coding of the REPEAT method are also included.

For the Ewald summations, the real space cutoff has been fixed to be $R_{cut} = 9$ Å, and the reciprocal space cutoff order has been fixed to be $k_{max} = 7$. The magnitude for the width of the Gaussian screening charge distribution was set as $\alpha = \sqrt{\pi}/R_{cut}$. These are typical values of the Ewald summation parameters recommended in classical molecular dynamics packages. We note that, when calculating the contributions of the real and reciprocal space series, the speed of convergence is not an issue because such series only have to be computed one time during the fitting process.

Only those grid points located outside the VDW spheres centered on each atomic site have been considered. Atomic regions due to the periodic images of the atoms are also excluded. The VDW radii used for this work were taken from the universal force field (UFF).[35] The grid points used for the fitting procedure coincided with the regular numerical FFT grid used in the periodic DFT calculation unless otherwise specified. The majority of the REPEAT charges reported in this work are unrestrained. In other words, the last term of the REPEAT error functional of eq 7 is set to zero. For results in which restraints were utilized, the electronegativity, χ, and self-Coulomb interaction parameters, J^{00}, of ref 34 are used.

Periodic DFT calculations were performed with several packages, VASP,[36] CPMD,[37] Gaussian 03,[38] and SIESTA.[39] All DFT calculations used the Perdew−Burke−Ernzerhof (PBE)[40] exchange and correlation functional. For the VASP calculations, the PAW method of Blöchl[41,42] was utilized with a plane wave cutoff of 600 eV. For the CPMD calculations the Goedecker−Teter−Hutter pseudopotentials[43] were used with a plane wave cutoff of 80 Ryd (∼1088 eV). The Brillouin zone integrations were performed using a 4 × 4 × 4 Monkhorst−Pack grid for the calculations of sodalite, natrolite, and zinc silicate, while a 14 × 14 × 14 sampling of the Brillouin zone was used for SnO_2, ZnO, and CdTe. For calculations of the isolated molecules and IRMOF-1, only the Γ-point was evaluated. True molecular calculations and CHELPG charges were performed with the Gaussian 03 package. RESP charges were generated using the AMBER 8 package.[44] To determine atomic charges using the Gasteiger's electronegativity equalization method[45] and Rappé's QEq method,[34] the Materials Studio modeling package version 4.1 was used. Except for the calculation of the isolated molecules, experimental geometries were used in evaluating the electrostatic potential.

We have implemented the REPEAT method as a FORTRAN code that is freely available upon request from the authors. The program reads a tabulation of the electrostatic potential in the Gaussian cube file format and can handle any unit cell shape.

To compare the electrostatic potential resulting from the fitted charges to that generated from the DFT calculation, we use the relative root-mean-square (RRMS) error defined in eq 17, where the sum runs over valid grid points or those lying outside the van der Waals radii of the atoms.

$$\mathrm{RRMS} = \sqrt{\frac{\sum_{\mathrm{grid}} \left(\phi_{\mathrm{QM}}(\vec{r}_{\mathrm{grid}}) - \phi_q(\vec{r}_{\mathrm{grid}})\right)^2}{\sum_{\mathrm{grid}} \phi_{\mathrm{QM}}(\vec{r}_{\mathrm{grid}})^2}} \tag{17}$$

IV. Results and Discussion

1. Charges Derived From Different DFT Packages. As a first validation of the REPEAT method, we first compare the charges derived from the electrostatic potential of sodalite depicted in Figure 1 using the conventional error functional of eq 2 and the charges derived using the REPEAT method with its modified error functional given in eq 4 (without the restraints).

Sodalite is a zeolite whose silicate framework is composed of tetrahedrally coordinated Si atoms linked through two coordinated oxygen atoms. This material has a cubic structure with a lattice parameter $a = 8.965$ Å and belongs to the $P\bar{4}3n$ space group. There are 36 atoms (12 Si, 24 O) in the unit cell, with only one symmetry unique Si and O atom. Three periodic DFT packages VASP, CPMD, and SIESTA were used to compute the electrostatic potential within the unit cell on a regular grid, using the PBE exchange correlation functional. The tabulation of the potential as a function of the square distance to the unit cell origin is shown in Figure 1, as already mentioned. The grid point density from each of the DFT calculations varied, but the total number of grid points available to perform the fitting (those outside of the VDW radii) was on the order of 3×10^5 for all cases.

Given in Table 1 are the ESP derived charges using the conventional error function of eq 2. The charges fluctuate widely depending on the DFT package used, and in some cases, the charges lack chemical meaning. It is notable that the RRMS errors also vary widely reaching up to 43% in the case of SIESTA. In contrast, the REPEAT charges also given in Table 1 are rather insensitive to the DFT code,

Table 1. ESP Charges of Sodalite Derived from the Electrostatic Potential Generated from Various Periodic DFT Packages Using the Standard Error Function and with The REPEAT method

	conventional ESP charges			REPEAT ESP charges		
atom	CPMD	VASP	SIESTA	CPMD	VASP	SIESTA
Si	+3.075	+10.522	+0.503	+1.355	+1.389	+1.151
O	−1.537	−5.261	−0.251	−0.677	−0.694	−0.575
RRMS error	0.193	0.289	0.426	0.066	0.066	0.064

although, small variations are to be expected since the three codes are using different types of basis sets and pseudopotentials. Table 1 not only shows that there are only small variations in the charges from one code to another, but it also shows that the root-mean-square errors remain quite similar and significantly smaller than those computed with the conventional functional.

To complement the numerical example above, it is also of value to consider a straightforward mind experiment to test the robustness of the REPEAT approach. In this experiment, it is assumed the existence of an ideal neutral periodic point charge distribution for which all the charges are known a priori. However, the tabulation of its corresponding electrostatic potential has been modified such that a small systematic constant error, δ_e, has been added to each one of its local values. If both the conventional as well as our REPEAT functional were to be applied on such a redefined potential, one can intuitively guess that the only way of recuperating the original values of the point charges is by using the REPEAT method. Moreover, the REPEAT method should provide correct charges regardless of the magnitude of the error level δ_e. Conversely, applying the conventional definition of the error functional makes the predictions for the charges level dependent and does not render the right values unless the error in the level is identically zero.

The REPEAT charges for sodalite can be compared to those derived for sodalite and related silicates using other methods where there is wide variation depending on the approach used. Nicholas and co-workers[46,47] developed charges of +1.1 e for Si and −0.55 e for O to model sodalite. These charges were derived from fragment calculations and adjusted to intrinsically account for cations in the sodalite framework when Si is replaced by Al atoms in the framework. For silicates, Si charges ranging from +0.4 to +1.91 e have been utilized.[47] Mulliken charges using the traditional atom centered basis functions gave a very wide range of charges for sodalite. Specifically, charges on Si were calculated to be +1.693, +1.723, and +0.965 e for the 3-21G, 6-31G, and 6-31G(d,p) basis sets, respectively. Using Gasteiger's electronegativity equalization method[45] gives charges of +0.624 e on Si and −0.312 e for O, while the QEq method of Rappé and Goddard[34] provides charges of +1.00 e on Si and −0.50 e on O.

2. Molecular Calculations. The previous results given in Table 1 reveal that the REPEAT algorithm minimizes the dependency of the charges with respect to the periodic DFT code. Although they show that similar charges are derived regardless of the arbitrary offset in the electrostatic potential, a natural question arises. Does the REPEAT method simply provide an optimum least-squares fit of the problem or do the charges generated represent more pragmatic charges similar in the spirit of existing ESP charge derivation methods? To examine this question, we study the REPEAT charges obtained from isolated molecule calculations and compare them to established molecular ESP charge derivation methods.

Table 2. Comparison between REPEAT, CHELPG, and RESP ESP Charges for Isolated Ammonia, Water, And Alanine Molecules

molecule	atom	CHELPG[a]	RESP[a]	REPEAT[b]	REPEAT (restrained)[c]
NH_3	N	−0.956	−0.752	−0.814	−0.813
	H	+0.319	+0.251	+0.271	+0.271
H_2O	O	−0.706	−0.704	−0.661	−0.661
	H	+0.353	+0.352	+0.330	+0.330
$CH_3NH_2CHCOOH$	C	−0.289	−0.321	−0.329	−0.318
	H	+0.075	+0.098	+0.091	+0.088
	H	+0.091	+0.098	+0.109	+0.107
	H	+0.083	+0.098	+0.098	+0.095
	N	−0.848	−0.795	−0.814	−0.805
	H	+0.334	+0.329	+0.322	+0.319
	H	+0.342	+0.329	+0.329	+0.327
	C	+0.328	+0.233	+0.264	+0.261
	H	+0.089	+0.022	+0.007	+0.006
	C	+0.530	+0.508	+0.612	+0.610
	O	−0.482	−0.464	−0.523	−0.523
	O	−0.462	−0.446	−0.481	−0.480
	H	+0.318	+0.311	+0.316	+0.314

[a] ESP charges derived from a molecular PBE/6-31G(d,p) calculation. [b] ESP charges derived from a VASP calculation of the molecules using cubic supercell of length 20 Å and the PBE functional. [c] Calculated with the restraining function introduced in eq 7.

Provided in Table 2 are the CHELPG and RESP charges of ammonia, water, and alanine resulting from a PBE/6-31G(d,p) isolated molecule calculation. Also presented in Table 2 are the REPEAT charges resulting from periodic VASP calculations of the molecules, using the same geometry as the molecular calculations. In the periodic calculations, a simple cubic simulation cell of $L = 20$ Å was used to minimize the interaction with the periodic images. Table 2 reveals that the CHELPG, RESP, and REPEAT charges are all in good agreement with one another. One should note that the CHELPG and RESP charges were derived from the same electrostatic potential, whereas the REPEAT charges were derived from a separate DFT calculation using a different basis set type and pseudopotential.[48] It is further notable that the total number of grid points (not shown in the table) used for the REPEAT charges was almost 2 orders of magnitude greater than those of the CHELPG and RESP methods.

Also provided in Table 2 are the REPEAT charges in which the quadratic restraints first introduced in eq 7 are applied. We have used a weighting factor defined as $w_j = 2a/J_j^{00}$ for each atom, where a is the weighting factor employed within the parabolic RESP approach. This is equivalent to using the weighting factor of 0.005 recommended by Baley et al. for the quadratic form of the RESP restraint.[18] The changes in the REPEAT charges resulting from the application of the restraints can be characterized as negligible. In these examples there are no embedded

atoms, and such negligible changes are desirable because the restraints will always act to increase the RRMS difference between the quantum mechanical electrostatic potential and the point charges. A detailed study of the most appropriate restraining parameters to evaluate REPEAT charges of periodic systems with embedded atoms will be the subject of an upcoming publication. In this paper, we focus on the general applicability of the REPEAT method to derive ESP charges for periodic systems but include the restraining formulation to retain the complete REPEAT formulation (eqs 7−16) in one place. It is worth mentioning that, in general, the charges predicted in molecular calculations are more susceptible to the RESP-type penalties than those obtained for periodic systems.

The results of Table 2 suggest that for isolated molecule calculations the REPEAT method does generate charges that are consistent with other ESP derived charge methods. Indeed, it can be shown that for the periodic calculation of isolated molecules the charges derived using the conventional ESP error functional converge to the REPEAT charges as the size of the simulation cell increases. If we rewrite the definition of the unrestrained REPEAT functional up to first-order corrections using eq 6 we obtain

$$F(\{q_j, \delta_\phi\}) = \sum_{\text{grid}} (\phi_{\text{QM}}(\vec{r}_{\text{grid}}) - \phi_q(\vec{r}_{\text{grid}}) - \Delta\phi_{q-\text{QM}} - I(\Delta\vec{p}_{q-\text{QM}}))^2 \quad (18)$$

where the new term $I(\vec{\Delta}p_{q\,-\,\text{QM}})$ is the integral over the dipole differences as appearing in the second term of eq 6. In the limit of an infinitely large box, the term $\Delta\phi_{q\,-\,\text{QM}}$ becomes the 'reference zero' of the quantum electrostatic potential because the 'reference zero' of the classical potential is strictly null. Therefore eq 18 can be rewritten as

$$F(\{q_j, \delta_\phi\}) = \sum_{\text{grid}} (\phi^*_{\text{QM}}(\vec{r}_{\text{grid}}) - \phi_q(\vec{r}_{\text{grid}}) - I(\Delta\vec{p}_{q-\text{QM}}))^2 \quad (19)$$

with $\phi^*_{\text{QM}} = \phi_{\text{QM}} - \Delta\phi_{q\,-\,\text{QM}}$ representing the 'true' quantum potential with respect to its well-defined zero. Thus, in order to prove that the REPEAT functional reduces to the conventional functional in the limit of an infinitely large box, we only need to show that $I(\vec{\Delta}p_{q-_{\text{QM}}})$ approaches zero in such limit. To achieve such a goal, we split the valid grid points into two regions of volume V_1 and V_2 as shown schematically in Figure 2. The first region is that lying outside of the VDW radii of all atoms and enclosed by a sphere centered at the charge origin of the molecule and with a radius R_1. We define R_1 as the distance from the charge origin to the furthest grid point lying on a VDW surface of an atom, where we have assumed that the charge origin is located at the center of a cubic simulation box. This first region of volume V_1 is that shaded in Figure 2. Next, the second volume is defined as that lying outside of the first sphere of radius R_1 but lying within a second sphere centered at the charge origin of radius $R_2 = \text{L}/2$, where L is the length of the simulation box. Using the above way of separating the grid into regions, the term $I(\vec{\Delta}p_{q-_{\text{QM}}})$ can be written as

$$I(\Delta\vec{p}_{q-\text{QM}}) = \frac{1}{V}\int \frac{(\vec{p}_{\text{QM}} - \vec{p}_q)\hat{r}}{|\vec{r}|^2} d^3r = \frac{1}{(V_1 + V_2)}\int_{V_1} \frac{(\Delta\vec{p}_{q-\text{QM}})\hat{r}}{|\vec{r}^2|} d^3r + \frac{1}{(V_1 + V_2)}\int_{V_2} \frac{(\Delta\vec{p}_{q-\text{QM}})\hat{r}}{|\vec{r}^2|} d^3r \quad (20)$$

The last integral in eq 20 turns out to be identically zero for symmetry reasons. Notice that the integration takes place on a region located between two spherical surfaces. Alternatively, the first integral yields a finite number, which is independent of the size of the simulation box. Nevertheless, having the first integral being multiplied by the inverse total volume allows us to make the contribution of the full $I(\vec{\Delta}p_{q\,-\,\text{QM}})$ as small as desired by increasing the simulation box size. In an analogous way, one can show that higher moment contributions could be made as small as desired. Thus, we have analytically shown that the REPEAT charges reduce to the standard ESP charges in the limit of an infinitely large simulation box.

To complement the previous analytical proof, we will numerically demonstrate that for molecular calculations the charges derived using the conventional error functional converge to the REPEAT charges as the size of the simulation cell increases. To do this, we have performed a series of water molecule calculations with VASP in which the cubic simulation cell is varied. At large box sizes, the water molecules can be considered noninteracting. However, as the box size, L, is decreased, the water molecule and its periodic images will begin to interact with one another. This situation is shown schematically in Figure 3. Since the dipole moments of the molecule and its images are aligned, the molecules should become more polarized as the simulation cell length is shortened. In other words, we would expect the fitted charges to become larger in magnitude as the length, L, is diminished.

Plotted in Figure 3 are the results of these calculations for both the REPEAT and the ESP charges derived by using the standard ESP error functional. As expected, the REPEAT charges become larger in magnitude as the simulation box size decreases. On the other hand, the ESP charges using the standard error functional become smaller in magnitude, in contradiction to what should physically occur. This is

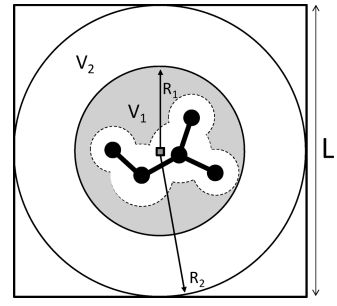

Figure 2. Schematic representation of an isolated molecule within a cubic simulation box of length L, defining the radii R_1 and R_2 of eq 20. The dotted outline represents the van der Waals radii of the atoms within the molecule. The shaded area represents V_1. The charge origin of the molecule is denoted by the square.

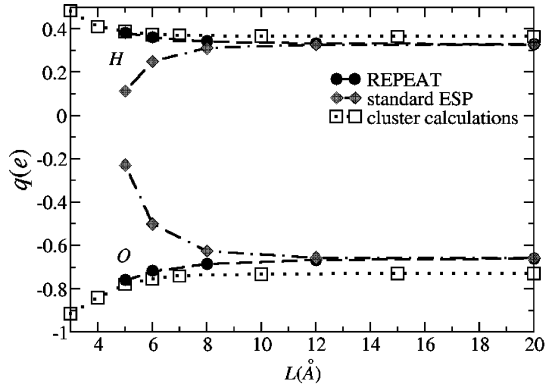

Figure 3. Charges of the O and H atoms of water as a function of the length of the simulation box using various methods.

undoubtedly a result of the ill-defined reference state in the periodic DFT calculations. As previously reasoned, Figure 3 shows the REPEAT charges and those derived using the standard functional converge to the same value as the simulation box increases. This suggests that the corrections due to the presence of the δ_ϕ term in eq 4 tend to zero as the simulation cell size increases.

To examine if the polarization of the REPEAT charges occurs in a physically suitable manner, we have performed a series of cluster calculations whereby a central water molecule is surrounded by 124 water molecules. The surrounding water molecules are positioned such that they correspond to the first and second shell of periodic images to mimic the truly periodic VASP calculations. The CHELPG charges of the central water molecule resulting from these cluster calculations are also plotted in Figure 3. At large simulation cell sizes, the charges resulting from the cluster calculation and the REPEAT charges converge to slightly different values. This is to be expected because of the differences in calculations, such as the nature of the basis sets used. Importantly, at short box lengths, the degree of polarization of the REPEAT charges is in good agreement with that observed in the cluster calculations. Thus, if molecular ESP derived charges are considered suitable for use in the simulation of molecular systems, then the results shown in Table 2 and Figure 3 suggest that REPEAT charges are equally appropriate for the simulation of periodic systems.

3. Dependency on Exclusion Radii and Grid Point Density. In this section, we examine how robust the REPEAT charges are with respect to the grid point density and the size of the VDW radii used to exclude grid points from the fitting procedure. To shed light on how the REPEAT charges vary with respect to these factors, we have selected two different porous materials: sodalite and zinc silicate. The results are summarized in Table 3 and Figure 4.

In most periodic electronic structure packages, the electrostatic potential is evaluated on a regular real space grid that is used for the discrete fast Fourier transform (FFT). For the practical application of the REPEAT method, it is of interest to know whether a typical FFT grid used in these calculations is appropriate for calculating the charges or whether a much higher grid density is required. To examine this, we have performed plane wave DFT calculations on sodalite and zinc silicate using an 80 Ry plane wave cutoff with the CPMD package. For sodalite, which has a simple cubic unit cell with length $a = 8.965$ Å, the plane wave cutoff results in a relatively dense $100 \times 100 \times 100$ grid point mesh. For zinc silicate, which has an orthorhombic unit cell with lengths $a = 10.07$, $b = 14.05$, and $c = 7.07$ Å, an 80 Ry plane wave cutoff in CPMD gives a default mesh of $120 \times 160 \times 80$ grid points. REPEAT charges of sodalite and zinc silicate have been evaluated using these default grid point meshes and are provided in Table 3 under the 'fine grid' subheading. The number of valid grid points per atom that lie outside of the VDW radii of the atoms that is used for the ESP fit is also given in Table 3. We then evaluated the REPEAT charges from the same DFT calculation, but where the grid points were uniformly trimmed. For example for sodalite, the grid points were trimmed by a factor of 8 and 125 to give $50 \times 50 \times 50$ and $20 \times 20 \times 20$ grid point meshes, respectively. These charges are given in Table 3 under the 'medium grid' and 'coarse grid' subheadings, respectively. Table 3 shows that the REPEAT charges remain essentially unchanged despite a drastic reduction in the number of grid points used in the fits. We note that the crystal symmetry was not imposed on the atomic charges during the fitting procedure. In the case of sodalite, which only has one symmetry unique Si atom and one unique O atom in a 36 atom unit cell, the standard deviation in the fitted charges was less than 2×10^{-6} e in all cases. For zinc silicate, the standard deviation was larger than with sodalite, but no greater than 6×10^{-3} e in all cases.

Table 3. REPEAT Charges for Sodalite and Zinc Silicate at Varying Grid Point Densities

		charge[a] or grid details		
molecule	atom	fine grid	medium grid	coarse grid
sodalite	Si	+1.355	+1.352	+1.372
	O	−0.677	−0.676	−0.686
	grid mesh	*100×100×100*	*50×50×50*	*20×20×20*
	grid points[b] per atom	*10 925*	*1 359*	*88*
zinc silicate[b]	Zn	+1.033(3)	+1.030(5)	+1.027(3)
	O	−0.493(3)	−0.494(3)	−0.485(3)
	Si	+1.387(3)	+1.383(4)	+1.377(6)
	grid mesh	*120×160×80*	*60×80×40*	*30×40×20*
	grid points[b] per atom	*6 215*	*778*	*98*

[a] Average charge (in e) over symmetry equivalent atoms. For zinc silicate, the standard deviation is reported in parentheses, while for sodalite, the standard deviation was 2×10^{-6} e or smaller in all cases. [b] Number of valid grid points outside of the VDW radii of the atoms used in the ESP fit. [b] Only the charges of selected Si and O atoms are presented.

For sodalite, a $20 \times 20 \times 20$ FFT mesh (the 'coarse grid') would be considered very sparse for any periodic DFT package we are familiar with and translates into only 88 fitting points per atom. Thus, we believe that the default FFT mesh settings used in most periodic electronic structure codes should be appropriate for deriving REPEAT charges.

Due to the lack of a unique definition of VDW radii, there should ideally be no significant oscillations in the REPEAT charges with modest variations in the radii used to exclude grid points from the fitting procedure. Shown in Figure 4 are the REPEAT charges of sodalite and zinc silicate plotted as a function of the factor used to uniformly scale the default VDW radii. In the scaling range of 0.8−1.3, the charges on

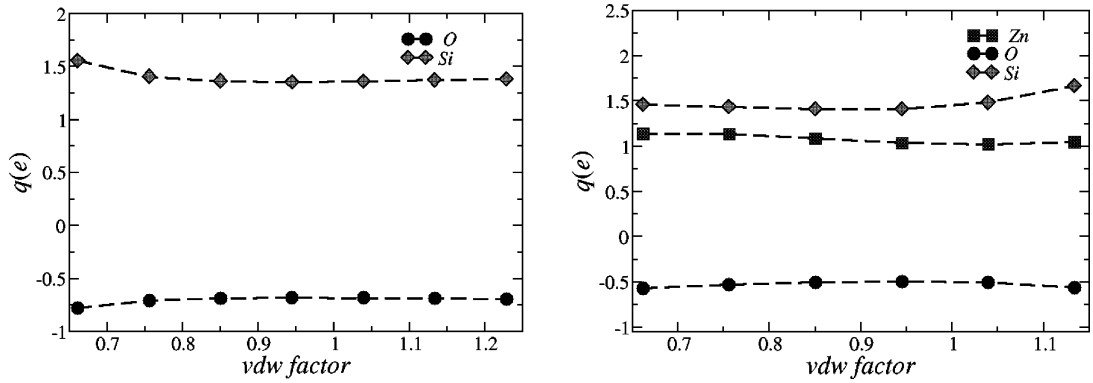

Figure 4. REPEAT charges for sodalite and zinc silicate under proportional scaling of the VDW radii used in the fitting procedure.

Si and O only differ by a maximum of 3.7% from the values derived using the default VDW radii. On an absolute scale, the O charges range from −0.703 to −0.678 e and for Si they range from +1.406 to +1.356 e. It is only when the VDW radii is rescaled by a factor of 0.7 that the charges begin to significantly deviate, differing by 14.7%. The REPEAT charges of zinc silicate are somewhat more sensitive to the VDW radii, with the charges deviating by more than 10% at scaling factors of 1.2 or greater and 0.7 or less. Such variations in the ESP fitted charges as a function of the VDW radii are related, in part, to alterations in the spatial geometry of the grid.

Our motivation for developing the REPEAT method was to generate charges for use in the classical simulation of microporous materials. Thus, we chose sodalite and zinc silicate to evaluate the method's sensitivity toward the VDW radii. We found that the total number of valid grid points could be drastically reduced without having a significant impact on the charges derived. As a result, we were curious as to how the method would perform when applied to densely packed solids, where the number of valid grid points available to perform the ESP fit is relatively small compared to microporous materials. For this purpose, we have examined the charges in zinc oxide, tin oxide, and cadmium telluride. Metal oxides like ZnO and SnO_2 enjoy a wide range of applications, ranging from gas sensors to transparent conductors in display technologies, while cadmium telluride has found important applications in solar cell technologies. VASP calculations of ZnO, SnO_2, and CdTe have been performed to generate an electrostatic potential for each compound. Using our default values for the VDW radii in these densely packed crystals resulted in mesh resolutions of approximately 10^3-10^4 grid points. Such numbers represent a total of at least 1 order of magnitude smaller than the typical grid sizes for microporous compounds, which have 3×10^5 points forming the grid. The variations in the charges as a function of the VDW scaling factor for zinc oxide, tin oxide, and cadmium telluride are given in Figure 5.

The REPEAT charges of oxygen in ZnO and SnO_2 appear to converge to the same approximate value of −1.25 e as a function of the VDW scaling. This is interesting since oxygen has a formal charge of −2 in both oxides. However, for zinc oxide, the convergence occurs at scaling factors greater than

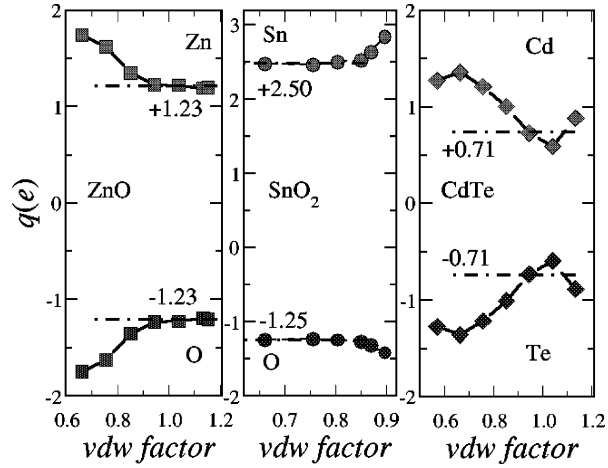

Figure 5. Comparison of the charges obtained for ZnO, SnO_2, and CdTe as a function of the VDW radii.

1, while it converges for values less than 1 for tin oxide. Another notable result is that the ESP charges for both the Zn and O atoms in ZnO are significantly greater in magnitude than those of the ESP charges of Zn and O in zinc silicate. This is consistent with the notion that the bonding in zinc silicate is more covalent, whereas the bonding in zinc oxide is considered more ionic in nature.

The results for CdTe are more difficult to analyze. As can be seen from Figure 5, the REPEAT charges oscillate and do not appear to converge as a function of the VDW radii used. This may be due to the fact that the charge is not as localized as in ZnO or SnO_2, whose bonding may be expected to be more ionic than in CdTe. This is supported by the fact that the charges in CdTe are notably smaller in magnitude than in ZnO, even though the formal charges are the same. Regardless of the underlying reasons for the behavior of the CdTe REPEAT charges, the results likely highlight the limitations of employing ESP derived charges for highly packaged systems.

4. Application to Porous Materials. Although one can foresee applications of the REPEAT method to various periodic systems, our motivation for developing the approach was our recent interest in simulating microporous and nanoporous materials, in particular MOFs. In this section, we apply the REPEAT method to two porous materials − a zeolite for which experimentally derived charges are available to compare with, and an MOF that has been the subject of

a number of recent computational studies where different partial atomic charges used in the classical simulations are also available to compare with.

The ultimate goal of any theoretical study with practical applications in mind is to be able to reproduce/predict experimental (or experimentally derived) data. Although atomic charges are not a physical observable, the detailed analysis of X-ray diffraction data can allow for electronic density modeling and estimation of atomic charges, for example, via multipolar refinement. Coppens and others have developed a few procedures to extract partial atomic charges from experimental X-ray crystallographic data in this way.[49−51] In the context of molecular simulation, Pearlman and Kim have compared experimentally derived charges to those typically used in force fields for a molecular crystal of a nucleic acid constituent, 2′-deoxycytidine-5′-monophosphate.[52] For microporous materials, one of the works that we are aware of where experimental partial atomic charges have been derived is that of Ghermani et al.,[53] who examined natrolite. Although the charges should not be strictly the same, we thought it would be of interest to compare REPEAT charges of natrolite using the experimental geometry to those derived experimentally.

Natrolite is a naturally occurring zeolite whose structure has been studied since the 1930s when Pauling first examined it.[53] Natrolite ($Na_2Al_2Si_3O_{10},2H_2O$) has a unit cell containing a total of 184 atoms and possesses an orthorhombic crystal with a space group symmetry *Fdd*2. We have evaluated the electrostatic potential from a VASP calculation using the experimental lattice parameters and the atomic positions.[53] The tabulation of the electrostatic potential had a resolution of $N = 200 \times 200 \times 72$ mesh points. In the original work in which natrolite charges were derived from X-ray diffraction data, its authors had fixed the atomic charge of the sodium ions to be unity (+1). Aside from basic chemical intuition, the authors justified the sodium charges by the sharpness of the 4s scattering factor of the cited element within the diffraction pattern. Therefore, the charges of all atomic sites were computed in a reference state in which the charge of the sodium atom was always +1.

Shown in Table 4 are the unrestrained REPEAT charges determined for natrolite generated from a VASP and a CPMD calculation. To compare to the experimentally derived charges, we have imposed a strong quadratic RESP-like penalty function on the sodium atoms to restrain those charges to +1. We have also derived REPEAT charges where the Na atoms were allowed to freely fluctuate. The REPEAT charges are in general agreement with the experimentally derived charges. Perhaps the most notable difference is that the relative magnitude of the REPEAT charges on Si_1 and Si_2 are the opposite of what was determined experimentally. Although the charge was restrained to match the experimental charge on Na, the overall agreement between the experimental charges and the REPEAT charges, as quantified by Δ^2, was slightly reduced compared to the unrestrained REPEAT charges. Since Mulliken charges are still widely used in simulation of microporous materials, these charges derived for natrolite are also provided in Table 4. These charges were obtained from a CPMD calculation whereby the plane wave orbitals were projected[54] onto a atom centered basis set composed of single-ζ Slater-type functions. The Mulliken charges are also in agreement with the experimentally derived charges, although the square difference, Δ^2, is approximately twice as large as the REPEAT charges. For reference, the REPEAT charges resulting from the CPMD calculation used to derive the Mulliken charges are also given in Table 4. Here it is notable that the REPEAT charges from CPMD and VASP (free Na charge) are very similar with a Δ^2 between one another of only 0.12.

Table 4. Atomic Charges Determined for Natrolite Using Various Methodologies

atom[a]	X-ray derived[b]	REPEAT (Na = +1),[d]	REPEAT[c] VASP	REPEAT CPMD	Mulliken[e]
Si_1	+1.84(12)	+1.376	+1.370	+1.304	+1.541
Si_2	+1.65(10)	+1.722	+1.516	+1.451	+1.484
Al	+1.51(11)	+1.616	+1.435	+1.383	+1.165
O_1	−0.90(5)	−0.898	−0.851	−0.814	−0.602
O_2	−1.21(5)	−1.057	−0.895	−0.862	−0.629
O_3	−1.03(5)	−1.107	−0.989	−0.954	−0.619
O_4	−1.07(5)	−1.099	−0.981	−0.946	−0.615
O_5	−0.87(5)	−0.767	−0.730	−0.697	−0.598
Na	+1.0000	+1.000	+0.882	+0.862	+0.930
O_w	−0.59(3)	−1.002	−0.858	−0.807	−0.488
H_1	+0.24(3)	+0.472	+0.408	+0.383	+0.318
H_2	+0.36(3)	+0.432	+0.377	+0.350	+0.331
Δ^{2f}	–	*0.501*	*0.489*	*0.609*	*1.156*

[a] Atom numbering is taken from ref 53. [b] Ref 53 where the charge of Na is set to +1.0. [c] REPEAT charges derived from a VASP calculation of natrolite using the experimental geometry of ref 53. [d] REPEAT charges derived from VASP where the charge of Na is restrained to +1.0. [e] Mulliken charge derived from the CPMD calculation in which the wave functions expanded in terms of plane waves are projected[54] onto a single-ζ atom centered Slater-type functions. [f] Δ^2 is the sum of the differences between the experimentally derived reference charges and the calculated charges squared, $\Delta^2 = \sum(q_i^{ref} - q_i^{calc})^2$.

One prototypical MOF material that has been the subject of several recent computational studies[12,23,32,55−59] is IRMOF-1, also known as MOF-5.[8] IRMOF-1 has attracted significant attention as the parent compound of materials that have potential hydrogen storage applications.[60] The structure of IRMOF-1 consists of inorganic Zn_4O clusters that are each connected to four 1,4-benzenedicarboxylate (BDC) organic linkers in an octahedral array to form a porous cubic framework. Figure 6 shows a ball and stick model of the unit cell of IRMOF-1, which consists of 424 atoms.

A number of groups have examined various aspects of IRMOF-1 using classical potentials where the atomic charges have been primarily generated from fragment ESP calculations. Using Monte Carlo simulations, Sagara et al.[32] examined the H_2 binding in IRMOF-1 using CHELPG charges derived from a molecular fragment calculation of $Zn_4O(CO_2CH_3)_5-BDC-Zn_4O(CO_2CH_3)_5$ at the PBE/6-31+G(d) level of theory. Schmid and co-workers[12] studied the diffusion of benzene in IRMOF-1 with charges derived from fragment calculations[58] in a manner similar to Sagara's work but using the Merz−Kollman[61] ESP charge derivation scheme. Greathouse and co-workers have examined the interaction of water in IRMOF-1 with molecular dynamics simulations using charges that are based on the CVFF force field[62] that were then manually adjusted to more closely match the charges of Sagara. The charges used in these

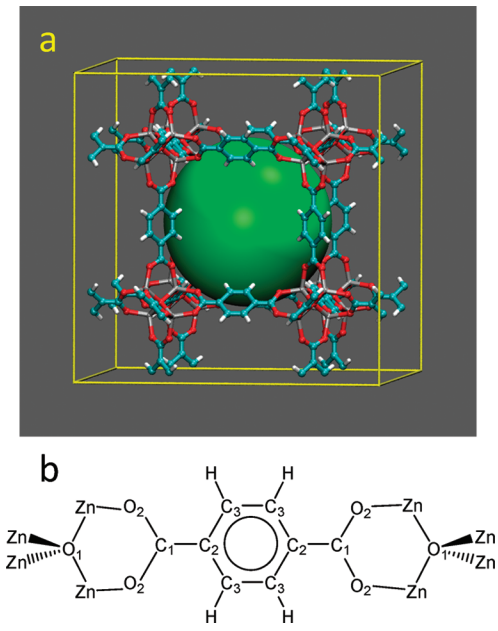

Figure 6. (a) Ball and stick representation of the unit cell of IRMOF-1 (blue = carbon, red = oxygen, white = hydrogen, gray = zinc). The green sphere represents the largest sphere that fits in the unit cell's cavity without touching the VDW's surface of the atoms. (b) Numbering of the unique atoms in IRMOF-1.

Table 5. Atomic Charges Determined for IRMOF-1

atom[a]	REPEAT	Sagara[b]	Schmid[c]	Greathouse[d]
Zn	+1.28	+1.31	+1.26	+1.20
O_1	−1.57	−1.79	−1.44	−1.20
O_2	−0.61	−0.63	−0.67	−0.60
C_1	+0.52	+0.62	+0.68	+0.60
C_2	+0.14	+0.05	+0.06	+0.00
C_3	−0.18	−0.12	−0.16	−0.10
H	+0.17	+0.12	+0.16	+0.10
Δ^{2e}	–	0.074	0.053	0.181

[a] Labeling of the atoms is given in Figure 6b. [b] Ref 32. [c] Ref 12. [d] Ref 56. [e] Δ^2 is defined in Table 2 with respect to the REPEAT charges.

simulation studies are given in Table 5 with the atom labeling provided in Figure 6b.

We have determined the REPEAT charges of IRMOF-1 resulting from the DFT calculation of the full periodic system using the CPMD package. These are given in Table 5 and show significant similarity to the charges reported by Sagara and Schmid resulting from the molecular fragment calculations and also to the charges of Greathouse that were manually adjusted to be similar to those derived by Sagara. These results further show that the REPEAT method produces sensible partial atomic charges for periodic systems. On the other hand, the results also put into question the need for an ESP charge derivation method for periodic systems, if molecular fragment calculations are adequate. We note that the suitability of ESP fragment calculations to ascertain partial charges in this particular compound relies on the fact that a chemically sensible and neutral moiety can be easily extracted from the framework. However, the extraction of fragments from a periodic solid can sometimes have nontrivial complications. For example, the need to 'cap' the fragment to satisfy unfilled valences can, among other things, compete with the ideal of a charge neutral fragment. One such example within the MOF family occurs with zeolitic imidazolate framework (ZIF) materials that have recently been shown to have remarkably high CO_2 storage capabilities.[5,7] ZIFs are composed of Zn^{2+} ions tetrahedrally coordinated to imidazolate anions, Im^-, in a 1:2 ratio to form structural topologies identical to those of zeolites. Extraction of chemically sensible fragment with one Zn ion, $[ZnIm_4]^{2-}$, leaves the imidazolate anions under coordinated. Adding another layer of Zn ions gives $[Zn_5Im_4]^{6+}$, unfortunately, a rather large net fragment charge. Alternatively, the REPEAT approach offers a simple yet robust approach to deriving atomic charges in periodic systems such that the complications associated with extracting a sensible molecular fragment do not have to be contended with because they are determined from a true periodic calculation. This means that minimal user intervention is required to generate the REPEAT charges, which is necessary for the automation of the charge derivation procedure. Such automation would be very useful in high throughput computational studies of nanoporous materials.

V. Conclusions

A simple and robust method to derive ESP charges in periodic systems is introduced that we have termed the REPEAT method. To the best of our knowledge, the REPEAT method is the first ESP fitting technique that properly circumvents the fundamental issue of the ill-defined offset in the electrostatic potential within periodic electronic structure calculations. By modifying the conventional error functional used in molecular ESP charge derivation methods, we have developed an approach that can be equally applied to molecular systems as well as to periodic systems such as zeolites. A formal analytical proof is provided showing that the conventional error functional used in molecular ESP calculations is a limiting case of the REPEAT method for an infinitely large simulation box. The methodology also implicitly includes corrections to the dipole and the higher multipole moments when evaluating charges of isolated molecules with a finite-size periodic simulation cell.

Validation of the new methodology is demonstrated with a variety of tests. First it is shown that when the conventional ESP functional is applied to a continuous periodic solid, the fitted charges obtained with different periodic electronic structure packages differ drastically. In some cases, the charges obtained are 'unphysical', for example +10.5 e for Si in sodalite. In contrast, the REPEAT method generates charges that are both physically reasonable and very similar between different electronic structure codes. Tests on isolated molecules were also performed. The REPEAT charges, resulting from periodic calculations of 'isolated' molecules, were found to be in good agreement with the CHELPG and the RESP charges resulting from nonperiodic molecular DFT calculations.

The stability of the REPEAT charges, with respect to the VDW radii used to exclude fitting grid points, as well as the grid point density was examined. For microporous materials, such as sodalite, the REPEAT charges were found to be very

stable upon variations of the aforementioned parameters, in particular the grid point density. To examine the limitations of the REPEAT method, several nonporous ionic solids were studied, specifically SnO_2, CdTe, and ZnO. The REPEAT charges of these systems were found to be chemically intuitive, however they did vary significantly with the VDW radii in some cases.

Although other applications can be foreseen, the original motivation for developing the REPEAT method was to generate partial atomic charges for the simulation of nanoporous materials, such as zeolites and metal organic framework materials. To demonstrate the applicability of the method to these systems, REPEAT charges were generated for the zeolite natrolite and the MOF IRMOF-1. Partial atomic charges have been determined from experimental X-ray diffraction data for natrolite and were found to be in good agreement with the REPEAT charges, although a direct correspondence is not expected. The second material, IRMOF-1, has been the subject of a number of computational studies where charges have been derived from molecular ESP calculations on fragments extracted from the periodic framework. The REPEAT charges were in excellent agreement with these charges. Extracting chemically sensible fragments from a periodic solid to perform molecular ESP charge derivation calculations is not always trivial and minimally requires some user intervention or manual adjustment. Therefore, the REPEAT method offers a straightforward and robust approach to deriving ESP charges for the simulation periodic systems that is amenable to automation.

Acknowledgment. We thank NSERC of Canada and the Canada Research Chairs program for funding. We are also grateful to CFI, the Ontario Research Fund, and IBM Canada for providing computing resources. We would like to acknowledge Dr. Saman Alavi for useful discussions and help with the preparation of some figures.

References

(1) *Handbook of Zeolite Science and Technology* Auerbach, S. M.; Carrado, K. A.; Dutta, P. K., Eds.; Taylor & Francis, Inc.: New York, 2003.

(2) James, S. L. *Chem. Soc. Rev.* **2003**, *32*, 276.

(3) Yan, Y.; Lin, X.; Yang, S. H.; Blake, A. J.; Dailly, A.; Champness, N. R.; Hubberstey, P.; Schroder, M. *Chem. Commun.* **2009**, 1025.

(4) Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M. *Science* **2002**, *295*, 469.

(5) Banerjee, R.; Phan, A.; Wang, B.; Knobler, C.; Furukawa, H.; O'Keeffe, M.; Yaghi, O. M. *Science* **2008**, *319*, 939.

(6) Dinca, M.; Long, J. R. *Angew. Chem., Int. Ed.* **2008**, *47*, 6766.

(7) Banerjee, R.; Furukawa, H.; Britt, D.; Knobler, C.; O'Keeffe, M.; Yaghi, O. M. *J. Am. Chem. Soc.* **2009**, *131*, 3875.

(8) Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M. *Nature* **1999**, *402*, 276.

(9) Chae, H. K.; Siberio-Perez, D. Y.; Kim, J.; Go, Y.; Eddaoudi, M.; Matzger, A. J.; O'Keeffe, M.; Yaghi, O. M. *Nature* **2004**, *427*, 523.

(10) Smit, B.; Siepmann, J. I. *Science* **1994**, *264*, 1118.

(11) *Computer Modelling of Microporous Materials*; Catlow, C. R. A.; Santen, R. A. v.; Smit, B., Eds.; Elsevier Academic Press: Amsterdam, The Netherlands, 2004.

(12) Amirjalayer, S.; Tafipolsky, M.; Schmid, R. *Angew. Chem., Int. Ed.* **2007**, *46*, 463.

(13) Salles, F.; Ghoufi, A.; Maurin, G.; Bell, R. G.; Mellot-Draznieks, C.; Ferey, G. *Angew. Chem., Int. Ed.* **2008**, *47*, 8487.

(14) Han, S. S.; Deng, W. Q.; Goddard, W. A. *Angew. Chem., Int. Ed.* **2007**, *46*, 6289.

(15) Han, S. S.; Goddard, W. A. *J. Am. Chem. Soc.* **2007**, *129*, 8422.

(16) Leach, A. R. *Molecular Modelling Principles and Applications*; 2nd ed.; Pearson Education Limited: Harlow, England, 2001, pp. 189−191.

(17) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.

(18) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.

(19) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300.

(20) Liu, D.; Zheng, C.; Yang, Q.; Zhong, C. *J. Phys. Chem. C* **2009**, *113*, 5004.

(21) Yang, Q. Y.; Zhong, C. L. *J. Phys. Chem. B* **2006**, *110*, 17776.

(22) Dubbeldam, D.; Walton, K. S.; Ellis, D. E.; Snurr, R. Q. *Angew. Chem., Int. Ed.* **2007**, *46*, 4496.

(23) Greathouse, J. A.; Allendorf, M. D. *J. Phys. Chem. C* **2008**, *112*, 5795.

(24) Blöchl, P. E. *J. Chem. Phys.* **1995**, *103*, 7422.

(25) Blöchl has developed a method24 to derive atomic partial charges from the total electron density of periodic plane wave calculations. However, the method is only valid for the calculation of isolated molecules or clusters when utilizing the supercell approximation in a periodic calculation.

(26) Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods* Cambridge University Press: Cambridge, U.K., 2004, 499−502.

(27) See Computational Details, Section III, for full details.

(28) Ewald, P. *Ann. Phys.* **1921**, *64*, 253.

(29) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. Lond. A* **1980**, *373*, 27.

(30) Smith, P. E.; Pettitt, B. M. *Comput. Phys. Commun.* **1995**, *91*, 339.

(31) *CPMD User's Manual* , Version 3.13.2, http://www.cpmd.org. Accessed August 6, 2009.

(32) Sagara, T.; Klassen, J.; Ganz, E. *J. Chem. Phys.* **2004**, *121*, 12543.

(33) Iczkowsky, R. P.; Margrave, J. L. *J. Am. Chem. Soc.* **1961**, *83*, 3547.

(34) Rappe, A. K.; Goddard, W. A. *J. Phys. Chem.* **1991**, *95*, 3358.

(35) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(36) Kresse, G.; Furthmüller, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 11169.

(37) *CPMD*, Version 3.13.2; IBM Research Division and MPI für Festkörperforschung Stuttgart: 2008.

(38) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; 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.; and Pople, J. A.; *Gaussian 03, Revision C.02*; Gaussian Inc.: Wallingford, CT, 2004.

(39) Soler, J. M.; Artacho, E.; Gale, J. D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez-Portal, D. *J. Phys.: Condens. Matter* **2002**, *14*, 2745.

(40) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(41) Blöchl, P. E. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50*, 17953.

(42) Kresse, G.; Joubert, D. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 1758.

(43) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 1703.

(44) Case, D. A.; Darden, T. A.; III, T. E. C.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, CA, 2004.

(45) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.

(46) Mabilia, M.; Pearlstein, R. A.; Hopfinger, A. J. *J. Am. Chem. Soc.* **1987**, *109*, 7960.

(47) Nicholas, J. B.; Hopfinger, A. J.; Trouw, F. R.; Iton, L. E. *J. Am. Chem. Soc.* **1991**, *113*, 4792.

(48) Although Gaussian 03 can perform periodic calculations, it cannot print out the electrostatic potential for periodic calculations.

(49) Hirshfeld, F. L. *Acta Crystallogr., Sect. B: Struct. Sci.* **1971**, *27*, 769.

(50) Hansen, N. K.; Coppens, P. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1978**, *34*, 909.

(51) Coppens, P.; Guru Row, T. N.; Leung, P.; Stevens, E. D.; Becker, P. J.; Yang, Y. W. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1979**, *35*, 63.

(52) Pearlman, D. A.; Kim, S.-H. *Biopolymers* **1985**, *24*, 327.

(53) Ghermani, N. E.; Lecomte, C.; Dusausoy, Y. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *53*, 5231.

(54) Sanchez-Portal, D.; Artacho, E.; Soler, J. M. *Solid State Commun.* **1995**, *95*, 685.

(55) Greathouse, J. A.; Allendorf, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 10678.

(56) Greathouse, J. A.; Allendorf, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 13312.

(57) Mulder, F. M.; Dingemans, T. J.; Wagemaker, M.; Kearley, G. *J. Chem. Phys.* **2005**, *317*, 113.

(58) Tafipolsky, M.; Amirjalayer, S.; Schmid, R. *J. Comput. Chem.* **2007**, *28*, 1169.

(59) Kim, D.; Kim, J.; Jung, D. H.; Lee, T. B.; Choi, S. B.; Yoon, J. H.; Kim, J.; Choi, K.; Choi, S.-H. *Catal. Today* **2007**, *120*, 317.

(60) Rosi, N. L.; Eckert, J.; Eddaoudi, M.; Vodak, D. T.; Kim, J.; O'Keeffe, M.; Yaghi, O. M. *Science* **2003**, *300*, 1127.

(61) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(62) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins: Struc. Funct. Genet.* **1988**, *4*, 31.

CT9003405