JCTC Journal of Chemical Theory and Computation

Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator

Victor M. Anisimov,[†] Guillaume Lamoureux,[‡] Igor V. Vorobyov,[†] Niu Huang,[†,||] Benoît Roux,[§] and Alexander D. MacKerell, Jr.*[,†]

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland, Département de physique, Université de Montréal, C.P. 6128, succ. centre-ville, Montréal, Québec, Canada H3C 3J7, and Department of Biochemistry, Weill Medical College of Cornell University, New York, New York

Received September 28, 2004

Abstract: A procedure to determine the electrostatic parameters has been developed for a polarizable empirical force field based on the classical Drude oscillator model. Atomic charges and polarizabilities for a given molecule of interest were derived from restrained fitting to quantum-mechanical electrostatic potentials (ESP) calculated at the B3LYP/ cc-pVDZ or B3LYP/aug-cc-pVDZ levels on grid points located on concentric Connolly surfaces. The determination of the atomic polarizabilities requires a series of perturbed ESP maps, each one representing the electronic response of the molecule in the presence of a background charge placed on Connolly surfaces primarily along chemical bonds and lone pairs. Reference values for the partial atomic charges were taken from the CHARMM27 additive all-atom force field, and those for the polarizabilities were based on adjusted Miller's ahp atomic polarizability values. The fitted values of atomic polarizabilities were scaled to reflect the reduced polarization expected for the condensed media and/or to correct for the systematic underestimation of experimental molecular polarizabilities by B3LYP calculations. Following correction of the polarizabilities, the atomic charges were adjusted to reproduce gas-phase dipole moments. The developed scheme has been tested on a set of small molecules representing functional moieties of nucleic acids. The derived electrostatic parameters have been successfully applied in a preliminary polarizable molecular dynamics simulation of a DNA octamer in a box of water with sodium counterions. Thus, this study confirms the feasibility of the use of a polarizable force field based on a classical Drude model for simulations of biomolecules in the condensed phase.

1. Introduction

Computer simulations based on empirical force fields are now a standard procedure to investigate biological phenomena.[1] Empirical force field calculations, due to the simplicity of the potential energy function, allow for atomic detail studies of biomolecules with explicit representation of the condensed phase environment to be performed. However, it is essential that the force field accurately reproduces the experimental regimen to ensure the quality of results of such calculations.

The majority of force fields consist of electrostatic, van der Waals, and bonding energy terms calculated in a pairwise additive fashion.[2] The induced polarization, which arises from a perturbation of the electronic structure of the molecular

* Corresponding author phone: (410)706-7442; fax: (410)706-5017; e-mail: amackere@rx.umaryland.edu. Corresponding author address: 20 Penn Street, Baltimore, MD 21201.
[†] University of Maryland.
[‡] Université de Montréal.
[§] Weill Medical College of Cornell University.
[||] Current address: Department of Biopharmaceutical Sciences, UCSF, San Francisco, CA.

10.1021/ct049930p CCC: $30.25
Published on Web 12/03/2004

species in response to the external electric field, is typically incorporated implicitly by using enhanced fixed partial atomic charges reflecting the average polarization taking place in the condensed phase. Despite the apparent simplification, additive empirical force fields have been remarkably successful in modeling complex molecular systems for the last two decades.[3] However, there are shortcomings in the additive model,[4] emphasizing the need to account for many-body induced polarization effects in an explicit way, motivating the development of polarizable force fields.

Current polarizable models can be classified into three major categories: point dipole, charge transfer, and classical Drude oscillator (or shell models); information about each of these models can be found elsewhere.[4,5] While a variety of efforts are ongoing to apply the point dipole and charge-transfer methods to biological systems,[6−27] the classical Drude oscillator approach has only seen minimal attention.[28−30] The Drude oscillator polarizability model was first introduced by Paul Drude in 1900 as a simple approach to describe the dispersive properties of materials.[31] The classical version of this model has been successfully used in statistical mechanical studies of condensed systems to treat electronic polarizability[5] and has been recently implemented[32] into the CHARMM program.[33,34]

Creation of a force field for biomolecular systems traditionally starts with the development of a water model. The polarizable SWM4-DP water model based on the classical Drude oscillator formalism has recently been presented.[30] It was parametrized to reproduce properties of liquid water under ambient conditions as well as some gas-phase properties such as the dipole moment and the interaction energy of the water dimer. In the SWM4-DP water model the polarizability of the oxygen atom, which in this case is equivalent to the molecular polarizability (since no Drude particles for hydrogen atoms were considered), was found to be 1.04 Å^3, which is 0.724 of the experimental gas-phase molecular polarizability of water, 1.44 Å^3.[35] Such reduced polarizability, which is essential to reproduce liquid-phase properties of water, including the dielectric constant, has been attributed to the energy cost of overlapping electron clouds in the condensed phase opposing induction.[30,36−41]

In the present work, steps toward systematic development of a polarizable force field based on the classical Drude oscillator model are presented. The determination of the electrostatic parameters is considered as the first step toward this goal. This effort, which includes fitting of partial atomic charges and atomic polarizabilities to a series of electrostatic potentials (ESP) around a molecule, each in the presence of an individual background charge, is the main focus of the present study. The theoretical background for the classical Drude oscillator model and the methodological details of the electrostatic parameter fitting are discussed in the next section. Validation of the theoretical level used in the derivation procedure is presented followed by a discussion of the selection of the reference values for charges and polarizabilities. Then, an example is given for the application of that procedure to the cytosine base, a model compound for nucleic acids. Finally, the validity of the proposed approach is illustrated via a condensed phase simulation of a DNA octamer using a preliminary classical Drude polarizable force field.

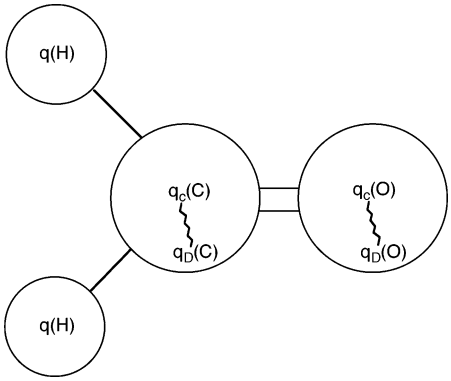

Figure 1. Classical Drude oscillator model using a formaldehyde molecule as an example. The displacements of the Drude particles, q_D, attached to the non-hydrogen atomic centers are exaggerated for clarity.

2. Theory and Methods

2.1. Classical Drude Oscillator Model. According to the classical Drude oscillator model, the polarizability is introduced by adding massless charged particles attached to each polarizable atom by a harmonic spring (Figure 1). Thus, a finite induced dipole is created and the partial atomic charge of atom A, $q(A)$, is redistributed between the Drude particle, $q_D(A)$, and the atomic core, $q_c(A)$. The positions of the Drude particles relative to the corresponding atomic centers are determined self-consistently by seeking the minimum energy consistent with the Born−Oppenheimer approximation. For the equilibrium position of the Drude particles, the atomic polarizability of atom A, $\alpha(A)$, is related to the charge $q_D(A)$ via the equation

$$\alpha(A) = q_D^2(A)/k_D \qquad (1)$$

where k_D is the force constant of the harmonic spring connecting a Drude particle to its corresponding atomic core. The magnitude of k_D is chosen to achieve small displacements of Drude particles from their corresponding atomic positions, $\mathbf{r}_D$, as required to remain close to the point-dipole approximation for the induced dipole associated with the atom−Drude pair.[32] Consequently, the atomic polarizability is determined by the amount of charge assigned to the Drude particle. Thus, the only adjustable parameters to be determined during parametrization of a polarizable atom A in the Drude model are the partial atomic charges $q(A) = q_c(A) + q_D(A)$. The development of a consistent protocol for the determination of $q_c(A)$ and $q_D(A)$ (or $\alpha(A)$) for a series of model compounds is, in part, the subject of the present study.

One of the major advantages of the classical Drude oscillator model is that it preserves the simple functional form of the pairwise additive force field and yet explicitly accounts for the electronic polarizability.[32] The electrostatic energy from the additive force field is substituted in the polarizable model by the Coulombic energy terms describing interactions between atomic cores and Drude particles and the self-energy of a polarizable atom treated via a harmonic term

$$U_{elec} = \sum_{A<B}^{N} \frac{q_c(A) \cdot q_c(B)}{|\mathbf{r}(A) - \mathbf{r}(B)|} + \sum_{A<B}^{N,N_D} \frac{q_D(A) \cdot q_c(B)}{|\mathbf{r}_D(A) - \mathbf{r}(B)|} + \sum_{A<B}^{N_D} \frac{q_D(A) \cdot q_D(B)}{|\mathbf{r}_D(A) - \mathbf{r}_D(B)|} + \frac{1}{2}\sum_{A}^{N_D} k_D |\mathbf{r}_D(A) - \mathbf{r}(A)|^2 \quad (2)$$

where N and N_D are the number of real atoms and Drude particles, respectively, q_c and q_D are the atomic core and Drude particle charges, respectively, and $\mathbf{r}$ and $\mathbf{r}_D$ are the positions of the real atoms and Drude particles, respectively. Modifying the classical force field energy function using eq 2 facilitates performing dynamical simulations with only minor modifications to existing programs, because the original energy functional form remains essentially unchanged. Computationally this means that the Drude particle positions for a given atomic configuration have to be self-consistently adjusted for each step of the dynamics simulation; however, such calculations are inefficient and rather expensive.[32] Therefore a molecular dynamics simulation algorithm based on the extended Lagrangian formalism[42,43] has been implemented, in which a small mass is attributed to the Drude particles, and the amplitude of the oscillators is controlled with a low temperature thermostat.[32] This technique allows the computationally expensive self-consistent field (SCF) regimen of molecular dynamics simulations to be avoided. Tests of the extended Lagrangian algorithm have shown that stable and accurate molecular dynamics trajectories can be generated, yielding liquid properties equivalent to the SCF regimen of molecular dynamics at a fraction of the computational cost.[32] Therefore, the classical Drude oscillator model for simulating atomic polarizability can be applied for molecular modeling in condensed media including macromolecular systems such as fully solvated nucleic acids, proteins, and lipid aggregates. However, the existing parameters from the nonpolarizable force fields need to be adjusted to take into account the presence of Drude particles.

2.2. Charge Fitting Scheme. The electrostatic properties of a molecular mechanics model with Drude polarizabilities are represented by atomic core charges $q_c(A)$ and Drude charges $q_D(A)$ producing the effective atomic charge $q(A)$ as their sum. Indeed, in the classical Drude oscillator polarizable model, the determination of atomic polarizabilities $\alpha(A)$ can be reduced to the determination of the partial charges of Drude particles, $q_D(A)$. Both $q(A)$ and $q_D(A)$ can be determined simultaneously, in a single fitting step.

Partial atomic charges are often obtained by optimizing the fit of an electrostatic potential ϕ^{MM} derived from the molecular mechanics (MM) model to a potential map ϕ^{QM} generated by quantum-mechanical (QM) calculations on a set of grid points $\{\mathbf{r}_g\}$ placed around the molecule. Although partial atomic charges of a nonpolarizable model can be extracted from a single QM potential map, adjusting the polarizabilities requires a series of response potential maps ϕ_p^{QM}, each one representing the altered charge distribution for the molecule in the presence of a small perturbing point charge z_p at a given position $\mathbf{r}_p$. A similar approach was used by Friesner et al. to derive parameters for the fluctuating charge and polarizable dipole models.[17,37,44] In our calculations the value of the perturbing charge z_p was arbitrarily chosen to be +0.5e. The MM potential at the g-th grid point (at the coordinate $\mathbf{r}_g$) for the molecule under the influence of a point-charge perturbation at position $\mathbf{r}_p$ is

$$\phi_{pg}^{MM} = \sum_{A}^{N}\left(\frac{q_c(A)}{|\mathbf{r}(A) - \mathbf{r}_g|} + \frac{q_D(A)}{|\mathbf{r}(A) + \mathbf{d}_p(A) - \mathbf{r}_g|}\right) + \frac{z_p}{|\mathbf{r}_p - \mathbf{r}_g|} \quad (3)$$

where $\mathbf{d}_p(A)$ is the Drude particle displacement from the corresponding atomic center position, $\mathbf{r}(A)$, in response to the perturbation p. The last term is the contribution from the perturbation charge itself.

During the fitting procedure, all core atomic and Drude charges have to be adjusted to minimize the discrepancy between the QM and MM potential maps, i.e., to minimize the following function:

$$\chi_\phi^2[q_c,q_D] = \sum_{p,g}(\phi_{pg}^{QM} - \phi_{pg}^{MM})^2 \quad (4)$$

The function ϕ_{pg}^{MM} has three unknown parameters: q_c, q_D, and $\mathbf{d}_p$. The first two are the subject of the standard least-squares fitting procedure, but the Drude particle displacement, $\mathbf{d}_p$, requires special consideration. Because of the implicit charge-dependence of the displacements $\mathbf{d}_p(A)$, the system of equations

$$\frac{\partial \chi_\phi^2}{\partial q(A)} = 0 \quad (5)$$

where $q(A)$ designates either q_c or q_D assigned to an atom A has to be solved iteratively. We use the Levenberg−Marquardt algorithm,[45] specially designed to minimize χ^2 functions (see below). First, Drude displacements, $\mathbf{d}_p$, are optimized to minimize total energy of the molecular system using atomic charges from the initial guess. This is followed by an ESP fitting step using the current positions of the atoms and Drude particles. The new set of fitted charges is again used to optimize the coordinates of the Drude particles. The iterative procedure is continued until eq 5 is satisfied.

Because the charge fitting problem is underdetermined, directly solving eq 5 usually leads to partial charges having poor chemical significance.[46] This is mainly due to the small contribution of some charges to the overall electrostatic potential associated with the screening of the charge on buried atoms by atoms located on the periphery of the molecule. To optimize individual charge contribution to the minimization function, it is necessary to penalize charge deviations from "chemically intuitive" reference values, as long as the penalty does not significantly deteriorate the quality of the fit. This requirement motivated the inclusion of restraints during charge fitting, referred to as restrained electrostatic potential (RESP) fitting. The original RESP scheme of Bayly et al.[46] minimizes

$$\chi^2 = \chi_f^2 + \chi_r^2 \quad (6)$$

through adding a penalty term in one of the two following forms

$$\chi_r^2 = w\sum_A^N [q(A) - \bar{q}(A)]^2 \quad (7)$$

$$\chi_r^2 = w\sum_A^N [\sqrt{q^2(A) + b^2} - b]^2 \quad (8)$$

where w is a weighting constant. The first restraint (7) is forcing the charges $q(A)$ to their "reference" values $\bar{q}(A)$, and the second restraint (8) favors smaller magnitude charges, where b is a hyperbolic stiffness parameter.

The RESP scheme can be generalized to the presence of Drude particles yielding the following parabolic (9) and hyperbolic (10) equations, respectively,

$$\chi_r^2 = \sum_A^N \{w[q(A) - \bar{q}(A)]^2 + w_D[q_D(A) - \bar{q}_D(A)]^2\} \quad (9)$$

$$\chi_r^2 = \sum_A^N \{w[\sqrt{q^2(A) + b^2} - b]^2 + w'_D[\sqrt{q_D^2(A) + b_D^2} - b_D]^2\} \quad (10)$$

where w and w_D are weighting factors for real atoms and Drude particles, respectively; b and b_D are the respective stiffness constants; $q(A)$ is the atomic charge representing the sum of atomic core and Drude particle charges, and $\bar{q}(A)$ is the reference charge. Due to the Drude charge − atomic polarizability formal equality (1) postulated by the classical Drude oscillator model, eqs 9 and 10 effectively lead to restraining the atomic polarizabilities.

To allow for additional flexibility of the fitted charges and polarizabilities, flat well potentials can be introduced into the fitting procedure. The parabolic restraint can be used only on the amount of charge deviating from $\bar{q}$ by more than a fixed charge tolerance q_{flat}. This allows the charge q to vary at no cost within the interval $[\bar{q} - q_{flat}, \bar{q} + q_{flat}]$ and creates a restraint only when the deviation is larger than q_{flat}. The charge fitting algorithm outlined above is implemented in the module FITCHARGE in the latest release of the CHARMM program.[33,34]

2.3. Grid Generation and Placement of Perturbation Charges. Electrostatic charge fitting procedures traditionally use QM electrostatic potentials determined on a cube based grid, with the grid points placed at an equidistant separation from each other. A limitation of this approach is the nonunique definition for the selection of the axes for the cube and lack of control over significance of each grid point. One approach to eliminate the orientation dependence is to use random point generation within the defined cube. However, this solution is not ideal due to reproducibility issues. In addition, for computational efficiency it seems reasonable to avoid placing grid points in regions having minimal chemical relevance. This is especially important when the molecular shape is significantly different from a spherical form. The well-known surface reflecting molecular shape, the Connolly surface,[47] is ideal for grid point placement. Such a structurally aware grid reduces the number of required grid points, facilitating the least-squares fitting by reducing the number of linear equations to solve, as well as maximizing the information content in selected regions around a molecule with the most chemical significance. The methodology is expected to be particularly advantageous for large molecules.

In practice, a Connolly surface is generated by overlapping atomic spheres and preserving only those regions of the spheres located on the periphery of a molecule. Surface building is based on using the van der Waals radii for the atomic spheres. Multiplication of all the atomic radii of the molecule by a constant creates a Connolly surface at the desired distance from the atomic centers. Choosing several different multiplication constants allows for creation of a set of nonintersecting Connolly surfaces. Similar considerations, as discussed above regarding the grid point placement, also apply to the placement of perturbation charges around the molecule. Multiple Connolly surfaces carrying the perturbation charges are required to probe molecular polarizability at different distances from the atomic centers. The chosen solution provides a simple mechanism to control the locations and number of point charges to place.

The simplest technique for point placement on a particular Connolly surface is to generate equidistant points on the atomic spheres at a selected density and to delete all the points which are within the overlapping spheres. The algorithm for placing grid points on the atomic sphere moves an atom to the center of the coordinate system, assigns the atom a unit radius, and generates points according to the sphere equation $x^2 + y^2 + z^2 = 1$. Next, the atomic sphere is scaled to the desired radius, and coordinates of the grid points are translated accordingly. Then a back translation of the atomic center to its original position is performed with simultaneous translation of coordinates of the generated grid points. These steps are applied to each atom of the molecule, and grid points in overlapping atomic spheres are deleted. This creates the final set of points situated on the Connolly surface.

By changing the number of grid points on a particular Connolly surface one can increase or decrease the contribution of that surface to the fitted molecular properties. Typically, large contributions from the nearest and most distant Connolly surfaces should be avoided; the nearest surface is approaching distances where the deviation from the atomic point charge approximation of the electronic distribution is still nonnegligible, while the contribution of the most distant surface may shift the accuracy of charge fitting to larger distances than those at which hydrogen bonding occurs. The following rationale may be used to identify the minimal number of Connolly surfaces for the placement of perturbation charges and grid points. The first layer of perturbation charges may be placed at distances typical for hydrogen bonds followed by a layer of grid points to detect an immediate ESP change caused by a perturbation charge. Several layers of perturbation charges are needed to adequately capture the orientational dependence of molecular polarizability. Therefore at least one more layer of perturbation charges is necessary followed by one more layer of grid points. Finally one additional distant layer of grid points is necessary to resolve the molecular dipole moment, which is a far-field molecular property. These operations yield five

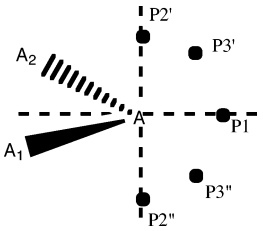

Figure 2. The placement of perturbation charges (P) for polarizability probing the region around lone pairs of sp^3 hybridized oxygen or sp^2 hybridized nitrogen atoms (A).

nonintersecting Connolly surfaces in total, two for perturbation charges and three for grid points, illustrated as (1) charges, (2) grid, (3) charges, (4) grid, and (5) grid according to increasing distance from the atomic centers. This number can be expanded as required based on computational tests.

The charge placement method employs a set of additional rules to the Connolly surface construction that further reduces the number of the charges. These additional rules are designed to assign perturbation charges to places of chemical significance, i.e., along covalently connected atoms and lone pairs where polarizability is expected to be largest. The perturbation points are prioritized into three groups according to the order of their generation. Charges placed along chemical bonds are first generated. If two atoms, A and B, are linked by a covalent bond, two perturbation points along the line A−B are generated that intersect with the corresponding Connolly surfaces on opposite sides of the bond. A newly generated charge will not be saved if it is too close to a previously generated charge. The distance criteria imposed in this study is 1.5 Å. The second group of perturbation points is created to sample regions around lone pairs on sp^3 hybridized oxygen or sp^2 hybridized nitrogen atoms, A, when these atoms have just two covalently bound neighbors (A1 and A2). In this case a bisector line is drawn in the plane of the covalent bonds of atom A, dividing the valence angle (A1-A-A2) in half (Figure 2). On the side of the lone pairs a perturbation point, P1, is generated along this line where it intersects with the Connolly surface. Two more points, P2′ and P2″, are generated on the line coming through point A, perpendicular to the plane defined by atoms A1, A, and A2 on the two opposite sides of the plane. Two other points, P3′ and P3″, are generated on the bisector line of angles P1-A-P2′ and P1-A-P2″. All these points are placed on the intersection of the Connolly surface with the corresponding line.

The last group of charges is created to close gaps between the previously placed perturbation points providing nearly equivalent coverage of the molecular shape by perturbation charges. The generation of the grids and the placement of perturbation charges are performed using the stand-alone program CGRID developed in our laboratory.

2.4. Molecular Dipole and Polarizability. The dipole moment and polarizability are important molecular properties which may be used to validate the quality of derived electrostatic parameters through comparison with experimental and/or QM values. To calculate molecular dipole moments, μ, the position of Drude particles must be optimized self-consistently. Then μ can be calculated from the sum over all charges using the following equation

$$\mu = \sum_{A}^{N} q_c(A)\mathbf{r}(A) + \sum_{B}^{N_D} q_D(A)\mathbf{r}_D(A) \tag{11}$$

where N and N_D are the number of real atoms and Drude particles, respectively, q_c and q_D are the atomic core and Drude particle charges, respectively, and $\mathbf{r}$ and $\mathbf{r}_D$ are the positions of the real atoms and Drude particles, respectively.

Polarizability is a measure of the response of a molecular system to an external electric field. Experimental measurement of isotropic molecular polarizability is conventionally conducted by measuring the refractive index η[48]

$$\alpha = \frac{3}{4\pi n}\frac{\eta^2 - 1}{\eta^2 + 2} \tag{12}$$

where n is the number density of the gas or liquid. In the present MM picture, the fast linear response of the electronic density to the excitation of a beam of light is modeled by the quasi-instantaneous readjustment of the Drude particles. The total molecular polarizability α of the MM model can be calculated in analogy with its standard QM definition, as the sum over all excitation modes of the square of the response dipole divided by the excitation energy. Being a linear-response coefficient, it can be summed over the vibration modes ν of the molecule obtained from normal-mode analysis of the polarizable MM model, complete with atomic cores and Drude particles. In terms of the Cartesian normal mode vectors $\mathbf{A}_\nu$, the components of the tensor α are

$$\alpha_{ij} = \sum_{\nu > \nu_0} \frac{\partial \mu_i}{\partial \mathbf{A}_\nu} \cdot \frac{\partial \mu_j}{\partial \mathbf{A}_\nu} \left(\frac{\partial U}{\partial \mathbf{A}_\nu} \right)^{-1} \tag{13}$$

where μ_i are components of the molecular dipole moment and U is the potential energy of the model. The isotropic polarizability values α_{iso} can be calculated as a trace of the molecular polarizability tensor, i.e., $\alpha_{iso} = 1/3(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$. This formula was implemented in the VIBRAN module of the CHARMM program. Only the very-high-frequency normal modes, attributed to Drude particle excitations, are summed, since the polarizability is modeled through the movement of these auxiliary particles. Given the much lighter mass of the Drude particles, a frequency cutoff at $\nu_0 = 5000$ cm^{-1} ensures that the normal modes associated with the vibrations of nuclei are excluded. The zero-frequency rotation and translation modes are ignored as well. It should be noted that the calculation of molecular polarizabilities requires full optimization of the molecular structure, since the harmonic approximation used in the vibrational analysis is valid only for equilibrium geometries.

2.5. Computational Details. The determination of electrostatic parameters for the polarizable force field is a multistep iterative procedure and employs both QM and empirical force field calculations. QM calculations were performed using the Gaussian 98 suite of programs.[49] Geometry optimizations were performed at the MP2(fc)/6-31G(d) level of theory for neutral species and at the MP2(fc)/6-31+G(d,p)

level for ions.[50−53] This level of theory provides molecular geometries consistent with available gas-phase experimental data, and it has been previously utilized during optimization of the CHARMM27 all-atom empirical force field for nucleic acids.[54] Geometry optimizations of adenine, cytosine, and guanine bases were conducted with the amino group planarity enforced, since it is believed that an amino group will acquire an approximately planar geometry due to hydrogen bonding with complementary bases and solvent molecules. No constraints were imposed during optimizations of other test molecules.

Selection of the appropriate level of theory for determination of the ESPs required for calculation of the charges and polarizabilities is important. Ideally, to ensure high accuracy in the ESP, a highly correlated method in conjunction with a very large basis set should be used, since computed dipole moments and molecular polarizabilities strongly depend on the level of theory and size of the basis set employed. However, when developing a force field for biomolecular systems it is necessary to have a large number of model compounds in the training set, with many of those compounds being relatively large, i.e., greater than 20 non-hydrogen atoms. All these requirements are unlikely to be satisfied in full and therefore a compromise between accuracy of the theoretical model and its computational performance is necessary.

QM calculations of the molecular electrostatic potentials were performed on MP2 optimized geometries using the B3LYP hybrid functional[55−57] and the correlation-consistent double-ζ Dunning cc-pVDZ and aug-cc-pVDZ basis sets.[58] Single-point energy B3LYP calculations were performed with the tight convergence criteria producing the target QM ESP maps. Cartesian coordinates for grid points and perturbation charges for the ESP calculation were generated by the program CGRID described above and read by Gaussian from external files. The generated QM ESP maps were extracted from the Gaussian output to use as input to the electrostatic parameter fitting by the FITCHARGE module of the CHARMM program. Restrained fitting using the RESP algorithm employed a penalty function with 10^{-5} Å^{-2} for the weighting factor. The flat well potential described above was applied allowing penalty-free deviation of charges by 0.1e in both directions from the corresponding reference values. Coordinates of the Drude particles were self-consistently adjusted after each change of the optimized charges during the charge fitting procedure to minimize the potential energy of the system, whereas coordinates of real atoms were fixed to the corresponding MP2 geometry.

Following the fitting procedure the atomic charges and polarizabilities were scaled. Polarizability scaling is necessary to reflect the reduced polarization expected for the condensed media and/or to correct for systematic underestimation of the experimental values by B3LYP/cc-pVDZ calculations (see below). Atomic charge scaling for neutral compounds can be performed to reproduce experimental or high-level QM gas-phase target molecular dipole moments. In addition, rounding of fitted and scaled charges and polarizabities to 3 decimal places was performed to facilitate their transferability while preserving the value of the net molecular charge. Technical details of the scaling and rounding procedure are described in the Supporting Information. In this study, similarly to the SWM4-DP water model, the polarization of only "heavy" (i.e. non-hydrogen) atoms is considered although the method can be easily extended to all atoms at the increase of the computational expense.

In all Drude polarizable CHARMM calculations, the Drude particles were attached to the real atoms via a harmonic spring with a force constant of a 1000 kcal/(mol·Å^2). This force constant is of sufficient magnitude to prevent large displacement of the Drude particle from its atom and thus ensure the validity of the point dipole approximation.[30] In this scenario, the atomic polarizabilities unequivocally determine the magnitude of Drude charges from eq 1. In addition, the sign of the charges on Drude particles is irrelevant due to the point dipole approximation. We chose q_D to be negative by analogy with the electron charge.

Molecular dynamics (MD) simulations were performed at 300 K and 1 atm pressure using the new velocity Verlet integrator[32] implemented in CHARMM.[33,34] The extended Lagrangian double-thermostat formalism[32] was used in all polarizable MD simulations where a mass of 0.1 amu was transferred from real atoms to the corresponding Drude particles. The amplitude of their oscillation was controlled with a separate low-temperature thermostat (at a $T = 1$ K) to ensure that their time course stays close to the SCF regimen.[32] A Nosé-Hoover thermostat with a relaxation time of 0.1 ps was applied to all real atoms to control the global temperature of the system. A modified Andersen-Hoover barostat with a relaxation time of 0.1 ps was used to maintain the system at constant pressure. Condensed-phase MD simulations were performed using periodic boundary conditions and SHAKE to constrain covalent bonds involving hydrogens.[59] A 1 fs time step was used for the extended Lagrangian polarizable MD simulation. The electrostatic interactions were treated using particle-mesh Ewald (PME) summation[60] with a coupling parameter 0.34 and 6th order spline for mesh interpolation. Nonbond pair lists were maintained out to 14 Å, and a real space cutoff of 12 Å was also used for the Lennard-Jones parameters within the atom-based force switch algorithm.[61] The long-range van der Waals correction recently implemented into CHARMM was also applied.[62,63]

Validation of the electrostatic parameters optimization method, along with proof of concept that the classical Drude oscillator model is applicable to biomolecular condensed phase simulations, was obtained via a MD simulation of DNA in solution. The preequilibrated GAGTACTC duplex DNA structure solvated in a box of water with sodium ions was taken from our previous study.[64] 1746 water molecules and 14 sodium ions were used. The solvated molecular system contains 9586 atoms including Drude particles. Starting from the CHARMM27 additive force field equilibrated system,[64] the Drude particles and then solvent molecules were minimized for 200 steps using the steepest descent (SD) algorithm with all DNA real atoms fixed. Then the minimized structure was subjected to a 20 ps NPT MD simulation with all DNA atoms, excluding the Drude particles, harmonically constrained with a mass weighted

Table 1. Summary of Experimental and QM Calculated Dipole Moments (Debye)[a]

			B3LYP/cc-pVDZ		B3LYP/aug-cc-pVDZ	
molecule		experimental		ratio		ratio
water	H_2O	1.855 ± 0.004	1.910	1.03	1.854	1.00
propane	C_3H_8	0.084 ± 0.001	0.070	0.83	0.096	1.15
isobutene	C_4H_{10}	0.132 ± 0.002	0.110	0.83	0.147	1.12
pentene-1	C_5H_{10}	0.500[b]	0.400	0.80	0.447	0.89
toluene	C_7H_8	0.375 ± 0.01	0.380	1.01	0.405	1.08
fluoromethane	CH_3F	1.858 ± 0.002	1.757	0.95	1.873	1.01
fluorobenzene	C_6H_5F	1.600 ± 0.08	1.394	0.87	1.603	1.00
chlorobenzene	C_6H_5Cl	1.690 ± 0.03	1.667	0.99	1.749	1.03
methanol	CH_4O	1.700 ± 0.02	1.576	0.93	1.680	0.99
ethanol (trans)	C_2H_6O	1.440 ± 0.03	1.483	1.03	1.598	1.11
ethanol (gauche)	C_2H_6O	1.680 ± 0.03	1.538	0.92	1.727	1.03
dimethyl ether	C_2H_6O	1.300 ± 0.01	1.200	0.92	1.306	1.00
tetrahydrofuran	C_4H_8O	1.750 ± 0.04	1.695	0.97	1.865	1.07
trimethylamine	C_3H_9N	0.612 ± 0.003	0.450	0.74	0.591	0.97
dimethyl sulfide	C_2H_6S	1.554 ± 0.004	1.471	0.95	1.620	1.04
ethanethiol (trans)	C_2H_6S	1.580 ± 0.08	1.557	0.99	1.669	1.06
acetaldehyde	C_2H_4O	2.750 ± 0.006	2.600	0.95	2.965	1.08
acetone	C_3H_6O	2.880 ± 0.03	2.789	0.97	3.167	1.10
acetic acid	$C_2H_4O_2$	1.700 ± 0.03	1.613	0.95	1.826	1.07
methylformate	$C_2H_4O_2$	1.770 ± 0.04	1.778	1.00	1.931	1.09
dimethylamine	C_2H_7N	1.010 ± 0.02	0.890	0.88	1.043	1.03
imidazole	$C_3H_4N_2$	3.800 ± 0.4	3.668	0.97	3.773	0.99
pyrazole	$C_3H_4N_2$	2.200 ± 0.01	2.217	1.01	2.316	1.05
pyridine	C_5H_5N	2.215 ± 0.01	2.050	0.93	2.299	1.04
trimethyl phosphate	$C_3H_9O_4P$	3.18[c]	3.475	1.09	3.723	1.17
			AVER	0.94		1.05
			SD	0.08		0.06

[a] Ratios are calculated with respect to experimental values. Data on water were not included in the calculation of the average ratio (AVER) and the standard deviation (SD). Calculations were performed using MP2(fc)/6-31G(d) geometries except for water, for which experimental gas-phase geometry was utilized. Experimental dipole moments are from ref 35. The experimental uncertainties are indicated where available. [b] Questionable results because of undetermined error sources. [c] Liquid-phase measurements, which may have large errors because of association effects.

force constant of 2 kcal/(mol·Å^2) to equilibrate the solvent around the DNA. The final structure from that simulation was then subjected to two 200 step SD minimizations, first for Drude particles and then for all atoms prior to initialization of the production trajectory. The simulation was run for 1000 ps with coordinates saved every 2 ps for analysis.

3. Results and Discussion

3.1. QM Calculations of Molecular Dipoles and Polarizabilities. The performance of B3LYP/cc-pVDZ and B3LYP/aug-cc-pVDZ calculations on MP2 optimized geometries in reproducing experimental gas-phase dipole moments and polarizabilities was first verified. Model compounds for which experimental values of molecular polarizabilities and/or dipole moments are available were chosen to represent different chemical classes as well as building blocks of biomolecules. The experimental and B3LYP calculated dipole moments are summarized in Table 1. The data indicate that both B3LYP/cc-pVDZ and B3LYP/aug-cc-pVDZ single-point energy calculations generally provide reasonable estimates of the dipole moment magnitudes. In many cases some underestimation of the experimental values can be noted for the B3LYP/cc-pVDZ calculations, whereas augmenting this basis set with diffuse functions in general results in a slight overestimation of the gas-phase molecular dipole moments. The average ratio of calculated to experimental dipole moments for compounds listed in Table 1, excluding water, was 0.94 and 1.05 for the cc-pVDZ and aug-cc-pVDZ basis sets, respectively.

Experimental and calculated molecular polarizabilities for selected compounds are summarized in Table 2. In all cases the experimental gas-phase molecular polarizabilities are underestimated by the B3LYP/cc-pVDZ calculations. The degree of the discrepancy is the largest for water (almost 50%) and substantially smaller for the remaining compounds. Importantly, for the majority of compounds the ratio of B3LYP to experimental polarizabilities is quite uniform, with an average value of 0.83 ± 0.06. As data in Table 2 demonstrate the substantial deviation of calculated from experimental values of polarizabilities can be corrected by augmenting the cc-pVDZ basis set by diffuse functions, which provides better quantitative agreement with the experimental data. Thus, B3LYP/ aug-cc-pVDZ QM calculations may be considered the method of choice for calculation of the ESP data for electrostatic parameter fitting. However, QM calculations with the aug-cc-pVDZ basis set are much more computationally expensive and become impractical for large and flexible molecules such as nucleosides. For example, for the guanine nucleoside cc-pVDZ and aug-cc-pVDZ basis sets consist of 331 and 554 basis functions,

Table 2. Summary of Experimental and Calculated Molecular Polarizabilities (Å^3)[a]

				B3LYP/cc-pVdz		B3LYP/aug-cc-pVdz	
molecule		experimental	ahp Miller		ratio		ratio
water	H_2O	1.45	1.41	0.78	0.54	1.39	0.96
ethane	C_2H_6	4.47	4.44	3.61	0.81	4.31	0.96
propane	C_3H_8	6.29	6.28	5.27	0.84	6.14	0.98
isobutane	C_4H_{10}	8.14	8.11	6.94	0.85	7.98	0.98
butane	C_4H_{10}	8.20	8.11	6.94	0.85	7.99	0.97
pentene-1	C_5H_{10}	9.65	9.76	8.42	0.87	9.88	1.02
pentene-2	C_5H_{10}	9.84	9.76	8.61	0.87	10.00	1.02
cyclohexane	C_6H_{12}	11.00	11.01	9.60	0.87	10.68	0.97
benzene	C_6H_6	10.00	10.43	8.57	0.86	10.35	1.04
toluene	C_7H_8	11.80	12.27	10.48	0.89	12.39	1.05
fluoromethane	CH_3F	2.97	2.52	1.92	0.64	2.53	0.85
fluorobenzene	C_6H_5F	10.30	10.34	8.57	0.83	10.35	1.00
chlorobenzene	C_6H_5Cl	12.25	12.36	10.20	0.83	12.51	1.02
methanol	CH_4O	3.32	3.25	2.41	0.73	3.19	0.96
ethanol	C_2H_6O	5.11	5.08	4.11	0.80	5.04	0.99
dimethyl ether	C_2H_6O	5.29	5.08	4.14	0.78	5.09	0.96
ethanethiol	C_2H_6S	7.41	7.44	5.70	0.77	7.30	0.99
acetaldehyde	C_2H_4O	4.59	4.53	3.69	0.80	4.58	1.00
acetone	C_3H_6O	6.39	6.37	5.28	0.83	6.37	1.00
acetic acid	C_2H4O_2	5.10	5.17	4.05	0.79	5.14	1.01
methylformate	$C_2H_4O_2$	5.05	5.17	4.06	0.80	5.12	1.01
pyrazole	$C_3H_4N_2$	7.23	7.72	5.76	0.80	7.32	1.01
pyridine	C_5H_5N	9.18	9.73	7.87	0.86	9.55	1.04
adenine	$C_5H_5N_5$	13.10	15.05	11.82	0.90	14.44	1.10
cytosine	$C_4H_5N_3O$	10.30	11.12	9.32	0.90	11.60	1.13
guanine	$C_5H_5N_5O$	13.60	15.68	12.56	0.92	15.40	1.13
thymine	$C_5H_6N_2O_2$	11.23	12.11	10.34	0.92	12.44	1.11
trimethyl phosphate	$C_3H_9O_4P$	10.86	10.86	9.27	0.85	11.23	1.03
				AVER	0.83		1.01
				SD	0.06		0.06

[a] Ratios are calculated with respect to experimental values. Data on water were not included in the calculation of the average ratio (AVER) and the standard deviation (SD). Calculations were performed using MP2(fc)/6-31G(d) geometries except for water, for which experimental gas-phase geometry was utilized. Experimental polarizabilities are from ref 35. When several experimental estimates were available, the most recent data were used. Calculated empirical polarizabilities are those obtained from atomic component values using the additive ahp scheme from ref 65.

respectively, and single-point energy B3LYP calculations using 1.5 GHz CPU take around 2 and 12 h, respectively. Approximately 100 such calculations are required for each conformation of the model compound to generate the perturbed maps of the ESP. Thus, for larger systems, B3LYP/cc-pVDZ calculations combined with the appropriate scaling, can be used for determination of ESP target data.

Based on the data in Tables 1 and 2, the following QM approach is suggested for determination of the electrostatic parameters. B3LYP/aug-cc-pVDZ calculations will be used for the calculation of QM response electrostatic potential maps. The scale factor 0.724 is applied to the fitted values of atomic polarizabilities to reflect the reduced polarizability required for the condensed phase simulations.[30] For larger molecules B3LYP/cc-pVDZ calculations are recommended. The underestimation of experimental polarizabilities by this level of theory must be corrected by applying the inverse of the average ratio of calculated to experimental gas-phase molecular polarizabilities 1/0.83 to the empirical values of $\alpha(A)$ as an additional scale factor. Combining this scale factor with the 0.724 factor yields an overall scaling factor of 0.87 that should be applied to the atomic polarizabilities after the fitting procedure. In addition, care must be taken when applying this level of theory due to the presence of outliers with respect to the calculated molecular polarizabilities (e.g. fluoromethane, Table 2).

3.2. Reference Values for Atomic Charges and Polarizabilities. Ideally, charges and polarizabilities can be determined via free fitting, as the resultant charges and polarizabilities are theoretically invariant to the initial guess. However, free fitting often results in physically unrealistic values of the charges[46] and polarizabilities. For instance, the polarizability tends to be attracted to one or a few atoms typically in the central region of a molecule and negative charges are often obtained for hydrogen atoms attached to aliphatic carbons from such fitting (see Table 3S of the Supporting Information). Therefore, the use of restraints is necessary to obtain chemically meaningful charges and polarizabilities. The task of careful selection of initial values for charges and polarizabilities becomes an important step since in restrained fitting these charges and polarizabilities are used as reference values in eq 9.

The classical Drude oscillator formalism employs the concept of atomic polarizability and assigns a corresponding

Table 3. Initial Values of Atomic Polarizabilities

atom (group)		polarizabilities (Å^3)	
symbol	Miller[a]	unscaled[b]	scaled[c]
$CH_3(sp^3)$	CTE+3H	2.222	1.844
$CH_2(sp^3)$	CTE+2H	1.835	1.523
$CH(sp^3)$	CTE+H	1.448	1.202
$C(sp^2)$	CTR	1.352	1.122
$CH(sp^2)$	CTR+H	1.739	1.443
$CH_2(sp^2)$	CTR+2H	2.126	1.765
$C(sp^2$, br)	CBR	1.896	1.574
−OH	OTE+H	1.024	0.850
−O−	OTE	0.637	0.529
=O	OTR4	0.569	0.472
$-O^{(-)}$	OTA[d]	0.858	0.712
$NH_2(sp^3)$	NTE+2H	1.738	1.443
$NH(sp^3)$	NTE+H	1.351	1.121
$N(sp^3)$	NTE	0.964	0.800
$NH_2(sp^2)$	NPI2+2H	1.864	1.547
$NH(sp^2)$	NPI2+H	1.477	1.226
$N:(sp^2)$	NTR2	1.030	0.855
P	PTE	2.063[e]	1.712

[a] Hybrid names used by Miller[65] are given in the second column along with the number of attached hydrogen atoms. [b] Unscaled united atom values of ahp polarizabilities can be used as reference values in the charge fitting to the B3LYP/aug-cc-pVDZ ESP potentials. [c] United atom values of ahp polarizabilities were scaled by the factor 0.83 reflecting the underestimation of experimental gas-phase polarizabilities by B3LYP/cc-pVDZ calculations. [d] The anionic oxygen atom type (OTA) was not present in the original Miller's ahp scheme and was added to reflect its higher polarizability compared to the ester oxygen (OTE). [e] The atomic polarizability of phosphorus atom was substantially changed from its original value in Miller's ahp scheme (1.538) to obtain a better estimate of the B3LYP molecular polarizabilities for phosphates. See details in the text.

unique parameter to individual atoms. Unlike molecular dipoles and polarizabilities, the atomic charges and polarizabilities of atoms in molecules are not well-defined quantities that can be unambiguously determined. A variety of different schemes have been proposed to obtain values of atomic polarizabilities using both QM and experimental molecular polarizabilities.[65−71] They can be categorized into element, group, bond, and hybrid polarizability schemes.[65] These methods can be also classified as additive, where the molecular polarizability is considered as a sum of atomic contributions, and nonadditive, which usually rely on the iterative solution of nonlinear equations to obtain molecular polarizability values. Nonadditive atomic polarizability schemes of Applequist[66] and Thole[67] have been commonly used to obtain polarizabilities in polarizable force fields employing point dipole induction models.[21,25,72] However, the atomic polarizability values from these as well as the majority of other nonadditive schemes can be considered as parameters of that model and, therefore, may be inappropriate for use in other polarization models. To the best of our knowledge, no atomic polarizability values for molecules other than water have been suggested for the classical Drude oscillator model.

Based on the classical Drude oscillator formalism, the atomic polarizability, $\alpha(A)$, of atom A is theoretically independent of the polarizability of the other atoms. However, in practice, atomic polarizabilities will be interdependent. Such interdependence is systematically minimized in force fields through assignment of atom types, which are expected to show transferability within a given class of chemical compounds. To facilitate transferability, the initial guess for atomic polarizability should also be taken from an additive polarizability scheme. The atomic hybrid polarizability (ahp) scheme of Miller[65] is a good example of such a scheme, and, after some modifications, it has been used in this work to provide an initial guess of atomic polarizabilities.

In the original Miller scheme the atomic polarizability contributions were obtained from least-squares fitting to experimental gas-phase molecular polarizabilities for approximately 400 organic compounds.[65] In addition, the Miller atomic hybrid polarizabilities α_A(ahp) depend not only on the identity but also the hybridization state of a particular atom, which is similar to an empirical force field atom type concept. According to the Miller scheme, molecular polarizabilities are obtained by summing up atomic hybrid polarizabilities α_A(ahp).[65] In most cases the sum of ahp polarizabilities is very close to experimental values of molecular polarizabilities (Table 2). Values of α_A(ahp) are available for most atomic hybrids which are encountered in biological compounds. Besides, the additivity of the model allows summing up atomic polarizabilities for functional groups and easy adjustment of α_A(ahp) values by applying scale factors (see below). Thus, the α_A(ahp) values were used as the initial estimates for atomic polarizabilities in the electrostatic parameter fitting procedure. It should be noted that molecular polarizability in the classical Drude oscillator model is not the additive sum of the atomic polarizabilities. However, α_A(ahp) values can still be used as an initial guess and then adjusted during the fitting procedure.

Since no polarizability is assigned to hydrogens in the current model, the initial guess for atomic polarizabilities was constructed by adding ahp polarizabilities of the heavy atoms and their covalently bound hydrogen atoms, thereby constituting the "united atom" approximation for atomic polarizability. If B3LYP/cc-pVDZ calculations are used for the calculation of the response ESP maps, these "united atom" atomic polarizabilities have to be scaled by the factor of 0.83 introduced above. Thus, their sum will reproduce the molecular polarizability at the B3LYP/cc-pVDZ level of theory, which will avoid biasing during the restrained fitting. Following the fitting procedure the scale factors described above are applied. No initial atomic polarizability scaling is required if the B3LYP/aug-cc-pVDZ level of theory is employed. Both unscaled and scaled atomic polarizabilities for biologically important atom types are given in Table 3.

A few small adjustments to Miller's polarizabilities were introduced to improve agreement with the QM results. A new hybrid atom type for anionic oxygen (OTA) was introduced to reflect the substantially higher electronegativity of anions and was derived based on the polarizability of the acetate anion. The atomic polarizability of the P atom was substantially increased to obtain a better estimate for different phosphate species, especially those with methyl groups attached to ester oxygen atoms. The comparison between B3LYP/aug-cc-pVDZ calculated and the sum of corrected ahp values from Table 3 for a few relevant compounds are

given in Table 1S of the Supporting Information. Overall, the satisfactory agreement between QM and empirical values of molecular polarizabilities was obtained for most compounds. It should be emphasized that these values represent initial estimates of atomic polarizability values, which are then adjusted in the fitting procedure. Since in the classical Drude oscillator model the atomic polarizability is directly related to a partial charge assigned to the Drude particle, the atomic and Drude charges are determined in one step through charge fitting to the series of perturbed ESP maps obtained from QM calculations.

It is also necessary to have a good choice of reference atomic charges used in eq 9. Three different initial guesses were considered, namely CHARMM27,[54] NBO,[73] and Mulliken[74] charges. Initial guesses for the polarizabilities were kept the same in all three cases. In general, it was found that the reference charge selection affects the final atomic charges but not the polarizability values. An example for such calculations for the cytosine base is given in Table 2S of the Supporting Information. It was also found that some molecular properties, such as residue-water interaction energies, were not very sensitive to the reference atomic charge scheme chosen (see also Table 2S). In addition, NBO and Mulliken atomic charges as initial guesses have some drawbacks. For instance, NBO charges on aliphatic hydrogens are often very large (0.20−0.30) even in nonpolar compounds such as alkanes. This results in the undesirable overpolarization of the C−H bonds. Mulliken charge assignment has a variety of shortcomings including a strong dependence on the basis set and unrealistic population assignment in the presence of diffuse and polarization functions, among others.[75] Moreover, charges obtained from QM models are conformation dependent, thereby complicating selection of the initial guesses. Therefore, the CHARMM27 charges were selected as the initial guess. This choice facilitates transfer of already existing internal and Lennard-Jones CHARMM27 force field parameters to the polarizable model, although optimization of these parameters will be necessary due to the new charge distribution and inclusion of explicit polarization. A disadvantage of using CHARMM27 charges as the initial guess is that they are limited to molecules for which the force field parametrization has been performed. Currently there is no simple scheme to derive these charges for new molecular residues although fitting procedures based on electronegativity equalization schemes[5] may be useful.

3.3. Atomic Charge and Polarizability Derivation for Model Compounds. The methodology described above has been applied for the derivation of the electrostatic parameters for the model compounds shown in Figure 3, which comprise the most important functional moieties of nucleic acids. Details of the fitting procedure will be discussed for the cytosine base. The geometry was optimized at the MP2/6-31G(d) level of theory with the amino group geometry constrained to be planar. Three Connolly surfaces for grid points and two surfaces for perturbation charges were generated (see Figure 4). The total number of grid points and perturbation ions was 1327 and 57, respectively. Surface parameters and the number of points placed on each surface

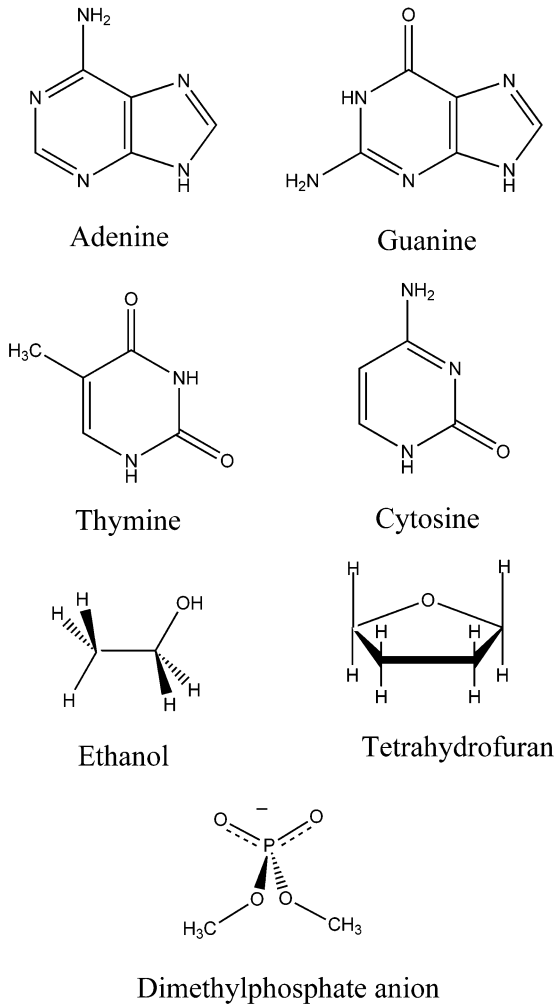

Figure 3. Model compounds used for the preliminary parameter development of the Drude polarizable CHARMM force field for nucleic acids.

are shown in Table 4. In total, 57 QM ESP maps were calculated, one for each placement of the perturbation charge, at the B3LYP/aug-cc-pVDZ level of theory. For other model compounds a similar strategy for the grid generation has been used. Reference partial atomic charges and atomic polarizabilities for the fitting procedure are shown in Table 4S of the Supporting Information. Fitting was performed under the RESP parabolic restraint with a 10^{-5} Å^{-2} weighting factor. Charges and polarizabilities were restrained separately to their corresponding initial values. The restraint was combined with a flat well potential with a half-width of 0.1e centered at the initial charge value. The restrained ESP fitting produced a final RMS error of $6.8 \cdot 10^{-4}$ e/Å with respect to the B3LYP potential, which is close to the RMS error of $3.9 \cdot 10^{-4}$ e/Å for the unrestrained fitting. The resultant atomic polarizabilities were then scaled by factor 0.724 to reflect reduced polarizability expected for the condensed phase (see above). This was followed by scaling the atomic charges to reproduce the B3LYP/aug-cc-pVZD gas-phase dipole moment. Final optimized partial atomic charges and atomic polarizabilities are shown in Table 4S of the Supporting Information for cytosine as well as other model compounds. In general, the fitted values of atomic charges decreased in

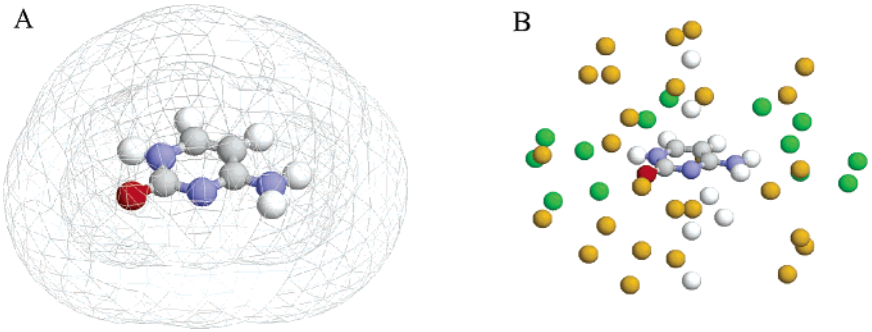

Figure 4. Electrostatic potential grid created based on Connolly surfaces around the cytosine base (A) and placement of perturbation charges around this molecule (B). For visualization purposes only grid 1 and grid 2 Connolly surfaces are displayed (see Table 4 for surface parameters). Perturbation charges are placed along chemical bonds (green spheres), around the N lone pair (white spheres), and in gaps between other charges (brown spheres) on the corresponding Connolly surfaces. The surfaces are not displayed in part B.

Table 4. Parameters Used for the Grid Point and Perturbation Charge Generation for the Cytosine Base

Connolly surface	size factor, f[a]	density factor, d	# of points
perturbation charges 1	2.2	1.1	46
grid 1	3.0	1.3	505
perturbation charges 2	4.0	0.1	11
grid 2	5.0	0.6	564
grid 3	6.0	0.2	258

[a] Size factors (multiplication constants) multiplied by the van der Waals radii of corresponding atoms (1.2 Å for H, 1.5 Å for C, 1.4 Å for N, 1.35 Å for O) define the distance from that atom to the generated Connolly surface. Density factors determine the relative density of points on a particular Connolly surface.

magnitude with respect to initial values. This is expected as polarization was taken into account implicitly in the CHARMM27 charges, whereas explicit polarization in the Drude polarizable CHARMM force field naturally results in decreased values of the atomic charge values in most cases (see Table 4S). In the fitting procedure redistribution of the polarizability also takes place. As a rule, the polarizability tends to accumulate on atoms that are located near the center of the molecules, which is mainly an artifact of the fitting procedure resulting from the inadequate sampling of the electrostatic potential for the "buried" central atom(s). The use of restraints allows a substantial reduction of this effect (see Table 3S).

In situations where model compounds have several stable conformers (e.g. dimethyl phosphate or ethanol), the grid generation and QM ESP calculations has been performed for the most relevant (i.e. the lowest-energy) nonequivalent conformers. For instance, gg and tg conformers have been used for the dimethyl phosphate electrostatic parameter derivation. The tt conformer of the DMP anion has not been taken into consideration since it lies substantially higher in energy (~3 kcal/mol compared to gg), and the corresponding conformation is not populated in nucleic acids. Then the fitting procedure was performed by simultaneously including the target ESP data for both conformers resulting in one set of charges and polarizabilities. The procedure tends to decrease the quality of the fit (i.e. increases the RMSE) but minimizes bias in electrostatic parameters toward one conformer.

The quality of the derived set of atomic charges and polarizabilities can be evaluated by the ability of the Drude polarizable model to reproduce the gas-phase dipole and polarizability. The corresponding data are presented in Table 5. The atomic charges and polarizabilities derived from the fitting procedure reproduce the reference experimental (if available) or QM gas-phase dipole moments within 3% for nucleic bases and tetrahydrofuran (THF) and within 10% for the stable conformers of the dimethyl phosphate anion (DMP) and ethanol, for which multiple conformers were used for fitting. The scaling of atomic polarizabilities results in the increase in the molecular dipole moments for the model compounds, which is corrected by the adjustment of partial atomic charges for all compounds except DMP. Since DMP is an ion the direct uniform charge scaling is not applicable. However the reproduction of the gas-phase dipole moment for DMP was considered to be not essential as the dipole moment of ionic species is undetermined, i.e., depends on the molecular orientation.

Ideally, the experimental and/or high-level QM gas-phase dipole moment values should be reproduced in Drude calculations for all neutral model compounds through the application of the atomic charge fitting. However, since the molecular dipole moment is also a function of the molecular geometry it may be necessary to perform fitting in an iterative fashion following adjustment of Lennard-Jones and internal parameters. In general, it should be noticed that the polarizable Drude model better reproduces gas-phase molecular dipole moments than additive CHARMM, which often overestimate them (Table 5) in order to reproduce condensed-phase properties of model compounds.

Concerning the molecular polarizabilities, those from the Drude model are underestimated by ca. 1−4% for nucleic bases, ~5% for ethanol and THF, and ca. ~14% for DMP compared to reference QM B3LYP/aug-cc-pVDZ data (Table 5). These differences may be a consequence of the different approaches used for determination of the QM and classical molecular polarizabilities as well as an artifact of the fitting procedure, which does not allow proper sampling for "buried" atoms. Future work will address the causes of these differences. However, the relative magnitudes of experimental and QM molecular polarizabilities are well reproduced by Drude polarizable CHARMM calculations.

Table 5. Summary of Calculated Molecular Dipole Moments and Polarizabilities for a Set of Model Compounds

					DRUDE	
model[a]	experiment	B3LYP/aug-cc-pVDZ	B3LYP/cc-pVDZ	CHARMM27[b]	unscaled[c]	scaled[d]
			Molecular Dipole Moments (Debye)			
ADE		2.43	2.35	2.94	2.43	2.25
CYT		6.72	6.24	7.88	6.70	6.92
GUA		6.97	6.84	7.61	6.96	6.81
THY		4.54	4.11	4.51	4.51	4.65
THF	1.75	1.87	1.70	2.34	1.91	1.72
DMP gg		5.43	4.88	*e*	5.42	*e*
DMP gt		4.82	4.31	*e*	4.75	*e*
DMP tt		2.70	2.35	*e*	2.43	*e*
ETOH t	1.44	1.60	1.48	2.36	1.70	1.68
ETOH g	1.68	1.73	1.54	2.40	1.81	1.73
			Isotropic Molecular Polarizability (Å^3)			
ADE	13.10	14.44	11.82		13.98	10.36
CYT	10.30	11.60	9.32		11.19	8.16
GUA	13.60	15.40	12.56		15.29	11.09
THY	11.23	12.44	10.34		12.09	8.72
THF		7.81	6.76		7.44	5.70
DMP gg		10.83	8.03		9.26	6.63
DMP gt		10.86	8.07		9.27	6.64
DMP tt		10.91	8.13		9.38	6.68
ETOH t	5.11	5.04	4.11		4.76	3.48
ETOH g		5.04	4.10		4.79	3.48

[a] ADE − adenine, CYT − cytosine, GUA − guanine, THY − thymine, THF − tetrahydrofuran, DMP − dimethyl phosphate anion, ETOH − ethanol. [b] CHARMM27 calculations were performed using the standard additive CHARMM27 force field for nucleic acids[54] and fully optimized geometry of the model compounds. [c] The fitted partial atomic charges and atomic polarizabilities were not scaled. MP2 optimized geometry was used, and only positions of Drude particles were optimized to make a direct assessment of the quality of the fit to the QM ESP. [d] The fitted atomic polarizabilities were scaled by a factor of 0.724. The full geometry optimization of model compounds was performed for molecular dipole and polarizability calculations. [e] Molecular dipole of ionic species depends on the molecular orientation and therefore cannot be compared directly for DMP using the QM and CHARMM optimized geometries.

3.4. Polarizable Condensed-Phase Molecular Dynamics Simulation of DNA. The ability to perform condensed-phase simulations of biologically relevant molecular systems is the ultimate goal of any macromolecular empirical force field. As a proof of concept of the Drude model and the presented electrostatic parameter derivation procedure, a MD simulation of the DNA duplex GAGTACTC in a SWM4-DP water box, including sodium counterions, was performed. Atomic charges and polarizabilities (Table 4S of the Supporting Information) of the model compounds shown in Figure 3 have been used. The values of atomic charges for terminal atoms had to be adjusted upon creating the covalent bonds in an oligonucleotide. For example, in pyrimidine bases the charge of the hydrogen H1 was summed into the N1 charge. The experimental polarizability of the Na^+ ion (0.157 Å^3)[76] scaled by 0.724 (see above) has been used.

Lennard-Jones (LJ) and bonded parameters for model compounds were also required for the simulation. Preliminary optimization of LJ parameters was performed to reproduce minimum interaction energies and geometries of model compounds with water and for base pairs. High-level ab initio LMP2/cc-pVQZ//MP2/6-31G(d) calculations for nucleic bases as well as THF and LMP2/aug-cc-pVTZ//MP2/6-31+G(d) calculations for the DMP were used as reference data.[77] In both the QM and empirical calculations the respective monomer geometries were fixed. The QM interaction distances were determined from the constrained MP2 optimization, and then interaction energies were calculated at the LMP2 level with a larger and more flexible basis set. In CHARMM the corresponding hydrogen bonding distances were sampled at a resolution of 0.01 Å, and the minimum interaction energies were obtained. The positions of Drude particles were self-consistently adjusted at every step of the potential energy scan. The hydrogen bond angles have been also sampled for some orientations of nucleic acid bases with water molecule, where it acts as a hydrogen bond donor to the carbonyl oxygen of the base.[54] The LJ parameters for model compounds were adjusted to minimize the difference between QM and CHARMM interaction energies and to reduce RMS deviation for both distances and energies across different complex orientations. Comparison of the present results with available CHARMM27 additive force field data[54] was also performed. The final values of the minimized base pair interaction energies are summarized in Table 5S of the Supporting Information, and those for the nucleic acid bases, DMP and THF with water are given in Tables 6S and 7S of the Supporting Information. For the nucleic acid bases the LJ parameter optimization was performed mainly based on the ab initio LMP2/cc-pVQZ//MP2/6-31G(d) base pairing interaction energies. The optimization of Lennard-Jones parameters for the DMP and THF was performed based on the LMP2//MP2 data on interactions with water.

Overall, the Drude model reproduces the QM base-base and model compound−water interactions at a level similar to that of CHARMM27. Base−base interactions are generally

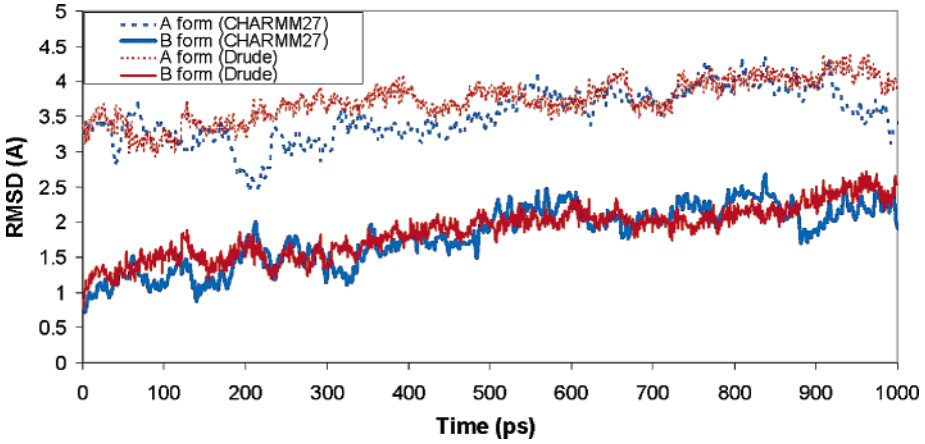

Figure 5. Root-mean-square deviation (RMSD) of heavy atom for six central residues of the GAGTACTC DNA duplex with respect to crystal structures of the A and B forms of DNA (dashed and solid lines) during the course of the MD simulation. RMSD from simulations using CHARMM27 pairwise additive force field and TIP3P water model (blue lines) and Drude polarizable CHARMM force field with SWM4-DP water model (red lines) are presented.

improved in the new model (Table 5S of the Supporting Information), although significant differences (e.g. the GG1 and GG2 interactions) that occur with CHARMM27 are still present. With the base−water interaction energies (Table 6S) the Drude model yields improved results for cytosine, while CHARMM27 is better for adenine, guanine, and thymine. Notably, the most significant improvement by the Drude model versus CHARMM27 is seen with dimethyl phosphate (Table 7S). This improvement includes the relative energies of the different interaction orientations and suggests that some of the largest gains in accuracy associated with the use of polarizable models may be expected with charged species. Thus, preliminary optimization of LJ parameters combined with electrostatic parameters determined using the methodology presented here yield a model that reproduces interaction energies to a level similar to CHARMM27 as judged by the reproduction of ab initio QM data. It is anticipated that more rigorous optimization of the LJ parameters, including the application of parameters obtained from condensed phase studies of small model compounds, will yield improved results for the studied interactions.

Selected internal parameters were also optimized. Equilibrium bond length and angle values were adjusted to reproduce average values of the bases targeting a survey of nucleic acid crystal structures from the NDB.[78] Force constants were adjusted to reproduce MP2/6-31G(d) frequencies scaled by 0.9434.[79] Torsional parameters for the C−O−P−O dihedral angles were adjusted to reproduce MP2/6-31+G(d) conformational energies for the DMP^- ion.[80]

Using this zero-generation polarizable model a MD simulation of the GAGTACTC DNA duplex in a box of SWM4-DP water box with sodium counterions was run for 1 ns. The results of this simulation were compared with a previously published MD simulation of this DNA sequence using the additive CHARMM27 force field and the TIP3P water model.[64] RMSD values for non-hydrogen atoms of the six central residues of the DNA duplex were calculated with respect to canonical A and B DNA structures for this duplex over the course of the simulation. These data are shown in Figure 5 and demonstrate that DNA structures from both Drude polarizable and CHARMM27 additive force field simulations remain closer to the B form versus the A form of DNA. More specific information may be obtained from the analysis of the base pairing interactions e.g. through monitoring of the N1···N3 distance between Watson−Crick bases. The average values over the course of the simulation are presented in Table 8S of the Supporting Information. These data indicate that the current Drude polarizable model gives reasonable agreement with the additive CHARMM27 force field. However, larger fluctuations of the N1···N3 distances for most base pairs from the Drude polarizable DNA simulation occur. Another test of the validity of the performed simulation is the analysis of the backbone dihedral parameters. Probability distributions of the backbone dihedral angles are presented in Figure 1S of the Supporting Information. The plots indicate that reasonable backbone dihedral angle probability distributions are obtained, being similar to those from the CHARMM27 MD simulation as well as data from a survey of the nucleic acid data bank (NDB)[78] in most cases. However, for some dihedrals there are noticeable differences in the relative conformer populations. The observed differences in the N1···N3 distances and in the dihedral distributions emphasize potential differences in molecular properties associated with the polarizable versus the additive CHARMM27 model. However, such differences also indicate the need for careful force field parametrization to be performed in order to properly implement a classical Drude based polarizable force field; such efforts are ongoing in our laboratory.

Importantly, the MD simulation was run in a reasonable amount of time. For instance, 50 ps of the extended Lagrangian DNA simulation took 23 CPU h to run on a 3 GHz Pentium IV computer. For comparison, the additive CHARMM27 force field simulation for the same system consumed approximately 5 h of CPU time. It should be noted that a larger time step of 2 fs was used in the CHARMM27 simulation, whereas a 1 fs time step was used for the extended Lagrangian simulation and that the TIP3P water model contains only 3 particles versus 5 particles in the SWM4-DP model.

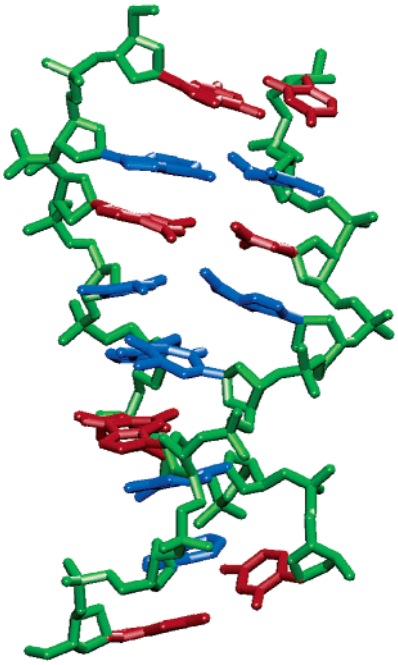

Figure 6. GAGTACTC duplex DNA molecule. Final structure from the 1000 ps Drude polarizable extended Lagrangian molecular dynamics simulation in a box of SWM4-DP water with sodium counterions.

Thus, applying the presented electrostatic parameter optimization methodology along with preliminary optimization of selected internal and LJ parameters produces a polarizable model of DNA that yields results comparable to those from the additive CHARMM27 nucleic acid force field. The final solvated DNA structure after 1000 ps is shown in Figure 6. In addition, the simulation system remained stable over the course of the trajectory as evidenced by a lack of significant drift in the molecular volume and potential energy of the system (not shown). These results indicate the applicability of the developed method for derivation of the electrostatic parameters and the resulting classical Drude based model of electronic polarizability for MD simulation studies of biological macromolecules in the condensed phase.

4. Conclusions

A general procedure for the determination of the electrostatic parameters for the classical Drude oscillator polarizable model, the partial atomic charges and atomic polarizabilities, is presented. This task is performed through fitting to a series of QM electrostatic potentials for a test molecule obtained in the presence of perturbation charges. Consequently, the partial atomic and Drude particle charges, where the latter are responsible for the atomic polarizabilities, are determined simultaneously in a single step. The QM level of theory for the electrostatic parameter fitting was determined via comparison of dipole moments and molecular polarizabilities for a variety of small organic molecules. The B3LYP/aug-cc-pVDZ level was selected based on this comparison, although the computationally less expensive B3LYP/cc-pVDZ level may be used for larger molecules with the appropriate scaling factors. Special emphasis was placed on the creation of the grid required for the electrostatic potential and placement of the perturbation charges. Versus the commonly used cubic grid, an approach based on the placement of grid points on a predetermined series of nonintersecting Connolly surfaces was developed. This approach reduces of the number of grid points by optimizing their placement around the test molecules. Perturbation charges were placed along chemical bonds, lone pairs, and in gaps between previous placed charges to provide equal coverage of the corresponding Connolly surfaces.

Consistent with previous work,[46] it was found that restraints were needed during the fitting procedure to avoid unphysical atomic charges and polarizabilities. Thus, generic reference values for the atomic charges and polarizabilities become important. Atomic charges from the CHARMM27 additive force field and atomic polarizabilities obtained using the atomic hybrid polarizability scheme of Miller[65] were identified as suitable reference values. However, it was necessary to adjust Miller's ahp atomic polarizability values to take into account the united-atom polarizability model used in this study, i.e., polarizabilities of H atoms are added to that of the corresponding heavy atom to which they are bonded. Furthermore, fitted values of atomic polarizabilities were scaled to reflect the reduced polarization which appears to take place in the condensed phase.[30,36−41] Studies on the SWM4-DP[30] and other polarizable models[36,37] have shown such scaling to be necessary to reproduce condensed-phase properties. In this study we used the same scaling factor as was used for the SWM4-DP water model. Studies are underway in our laboratory to determine if the scale factor based on the water molecule (0.724) is appropriate for other small organic molecules. These tests will involve the ability of the scaling approach to reproduce pure solvent properties, including dielectric constants and free energies of solvation. Use of scaling of the polarizabilities required readjustment of the partial atomic charges to ensure that gas-phase dipole moments were reproduced. The quality of the fits was evaluated via comparison of calculated molecular dipole moments and polarizabilities with the available gas-phase experimental and QM values.

The developed scheme for the determination of atomic charges and polarizabilities has been tested on a set of small molecules representing functional moieties of nucleic acids. All other parameters have been taken from the all-atom additive CHARMM27 force field for nucleic acids. Selected Lennard-Jones parameters were adjusted to reproduce QM data on interactions of model compounds with water as well as nucleic acid base pairing interactions. Selected internal parameters were optimized to reproduce experimental and QM data on molecular geometries, vibrational frequencies, and rotational barriers. The resulting zero-generation force field has been successfully applied in a 1 ns polarizable MD simulation of a DNA octamer in aqueous solution. This simulation validates the feasibility of the developed methodology for the determination of partial atomic charges and polarizabilities as well as the use of the Drude oscillator model to include electronic polarizability in biomolecular systems. Future efforts will apply the methodology developed in this work, along with an iterative parameter optimization scheme that includes the internal and LJ parameters, to

develop a nonadditive empirical force field for molecules of biological and pharmacological interest.

Acknowledgment. We gratefully acknowledge financial support from the NIH (GM51501) to A.D.M. and from the NSF (NSF 0415784) to G.L. and B.R. Computer time allocations were received from the Pittsburgh Supercomputing Center and the DOD ASC Major Shared Resource Computing and High Performance Computing.

Supporting Information Available: The comparison of QM and empirical molecular polarizabilities; initial and fitted values of atomic charges and polarizabilities for model compounds; the comparison of QM, CHARMM27, and Drude polarizable CHARMM calculations for nucleic base pairing interactions and interactions of model compounds with water; the comparison of CHARMM27 and Drude polarizable CHARMM average base pairing distances and backbone dihedral angle distributions from the DNA simulation; and the description of the technical details of the charge and polarizability scaling and rounding. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) MacKerell, A. D., Jr. Atomistic Models and Force Fields. In *Computational Biochemistry and Biophysics*; Becker, O. M., MacKerell, A. D., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; p 7.

(2) MacKerell, A. D., Jr. *J. Comput. Chem.* **2004**, *25*, 1584.

(3) Saiz, L.; Klein, M. L. *Acc. Chem. Res.* **2002**, *35*, 482.

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

(5) Rick, S. W.; Stuart, S. J. *Rev. Comput. Chem.* **2002**, *18*, 89.

(6) Sprik, M.; Klein, M. L. *J. Chem. Phys.* **1988**, *89*, 7556.

(7) Caldwell, J.; Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 9144.

(8) Wallqvist, A.; Berne, B. J. *J. Phys. Chem.* **1993**, *97*, 13841.

(9) Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 4180.

(10) Dang, L. X. *J. Phys. Chem. B* **1998**, *102*, 620.

(11) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672.

(12) Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. *J. Chem. Phys.* **1998**, *109*, 3077.

(13) Patel, S.; Brooks, C. L., III. *J. Comput. Chem*. **2004**, *25*, 1.

(14) Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1504.

(15) Chen, B.; Xing, J.; Siepmann, I. J. *J. Phys. Chem. B* **2000**, *104*, 2391.

(16) Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **2001**, *115*, 2237.

(17) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730.

(18) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(19) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671.

(20) Shelley, J. C.; Sprik, M.; Klein, M. L. *Langmuir* **1993**, *9*, 916.

(21) Gao, J.; Habibollazadeh, D.; Shao, L. *J. Phys. Chem.* **1995**, *99*, 16460.

(22) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 6208.

(23) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177.

(24) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386.

(25) Cieplak, P.; Caldwell, J. W.; Kollman, P. A. *J. Comput. Chem.* **2001**, *22*, 1048.

(26) Dang, L. X. *J. Phys. Chem. B* **1999**, *103*, 8195.

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

(28) Stuart, S. J.; Berne, B. J. *J. Phys. Chem.* **1996**, *100*, 11934.

(29) van Maaren, P. J.; van der Spoel, D. *J. Phys. Chem. B* **2001**, *105*, 2618.

(30) Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185.

(31) Drude, P.; Mann, C. R.; Millikan, R. A. *The theory of optics*; Longmans, Green, and Co.: New York [etc.], 1902.

(32) Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025.

(33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(34) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and Its Paramerization with an Overview of the Program. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, 1998; Vol. 1, p 271.

(35) *CRC Handbook Chemistry and Physics*, 84th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, 2003.

(36) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2004**, *120*, 9903.

(37) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2004**, *108*, 621.

(38) Morita, A. *J. Comput. Chem.* **2002**, *23*, 1466.

(39) Morita, A.; Kato, S. *J. Chem. Phys.* **1999**, *110*, 11987.

(40) in het Panhuis, M.; Popelier, P. L. A.; Munn, R. W.; Angyan, J. G. *J. Chem. Phys.* **2001**, *114*, 7951.

(41) Tu, Y. Q.; Laaksonen, A. *Chem. Phys. Lett.* **2000**, *329*, 283.

(42) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.

(43) van Belle, D.; Couplet, I.; Prevost, M.; Wodak, S. J. *J. Mol. Biol.* **1987**, *198*, 721.

(44) Banks, J. L.; Kaminski, G. A.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.

(45) Press: W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: Cambridge, 1988.

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

(47) Connolly, M. L. *Science* **1983**, *221*, 709.

(48) Bonin, K. D.; Kresin, V. V. *Electric-dipole polarizabilities of atoms, molecules, and clusters*; World Scientific: Singapore River Edge, NJ, 1997.

(49) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98; Gaussian, Inc.: Pittsburgh, PA, 1998.

(50) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(51) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.

(52) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta (Berlin)* **1973**, *28*, 213.

(53) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.

(54) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86.

(55) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(56) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

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

(58) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(59) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(60) Darden, T. A.; York, D.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.

(61) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667.

(62) Lague, P.; Pastor, R. W.; Brooks, B. R. *J. Phys. Chem. B* **2004**, *108*, 363.

(63) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.

(64) Pan, Y.; MacKerell, A. D., Jr. *Nucl. Acid Res.* **2003**, *31*, 7131.

(65) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.

(66) Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* **1972**, *94*, 2952.

(67) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(68) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399.

(69) Ewig, C. S.; Waldman, M.; Maple, J. R. *J. Phys. Chem. A* **2002**, *106*, 326.

(70) Stout, J. M.; Dykstra, C. E. *J. Phys. Chem. A* **1998**, *102*, 1576.

(71) Zhou, T.; Dykstra, C. E. *J. Phys. Chem. A* **2000**, *104*, 2204.

(72) Ding, Y.; Bernardo, D. N.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1995**, *99*, 11575.

(73) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.

(74) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.

(75) Leach, A. R. *Molecular Modelling: Principles and Applications*; Longman: Harlow, 1996.

(76) Mahan, G. D. *Phys. Rev. A* **1980**, *22*, 1780.

(77) Huang, N.; MacKerell, A. D., Jr. *J. Phys. Chem. B* **2002**, *106*, 7820.

(78) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751.

(79) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*; Gaussian, Inc: Pittsburgh, PA, 1996.

(80) Banavali, N. K.; MacKerell, A. D., Jr. *J. Am. Chem. Soc.* **2001**, *128*, 6747.

CT049930P