Atomic Charges for Large Molecules Derived from Electrostatic Potentials: Fragment Density Matrix Approach

Jung-Goo Lee and Richard A. Friesner*

Department of Chemistry, Columbia University, New York, New York 10027

Received: September 1, 1992

We describe an accurate and efficient approach for computing atomic charges of molecules built from density matrix elements of smaller molecules obtained from ab initio SCF calculations. The calculated atomic charges of $C_{10}H_{22}$, $C_9H_{19}NH_2$, $C_8H_{17}COOH$, and $C_9N_3O_3H_{17}$ are compared with those derived from conventional ab initio electrostatic potentials, and the results are uniformly of high accuracy.

Introduction

The representation of the electrostatic interactions of molecules via atomic point charges is at present the most commonly used approach in the construction of electrostatic force fields. A popular methodology is to compute these charges via quantum chemical calculations (ab initio SCF calculations,[1] semiempirical methods,[2] or density functional computations[3]) in which the charge density is obtained from the resultant wave function, the Coulomb field from this density is evaluated at the van der Waals radii of the atoms, and the point charges are adjusted to reproduce this Coulomb field using a least-squares fitting procedure. While there are of course fundamental questions about the implications of this electrostatic potential fit (ESP) approach in terms of the decomposition of the exact quantum mechanical energy, it has proven to be reasonably successful in the construction of molecular mechanics potentials which have then been used to reproduce molecular properties (e.g., solvation energies[4]).

Such an approach is computationally tractable for small molecules and, as computer power grows, to increasingly larger systems. However, for very large molecular assemblies the cost becomes prohibitive, particularly if one wishes to use higher quality quantum chemical methods with large basis sets.

The solution to this problem which is generally adopted is to calculate charges for molecular fragments and then build the charges for the large assembly from the fragment values. In many cases, this is done as an uncontrolled approximation. At the very least, it would be useful to have a more accurate method of generating charges for large systems to test the validity of the standard sort of fragment approaches.

The objective of this paper is to define such a method and to show on a series of test cases that it reproduces the charges obtained by quantum chemical computations for the entire molecule at a much lower cost. The method is straightforward modification of standard fragment methods, with one additional step which has a computational cost that scales similarly (as a function of system size) to quantum chemical computations on the entire molecule but a much smaller prefactor and lower memory requirements. It should therefore be tractable on systems containing hundreds of atoms.

The basis idea is to calculate *density matrix elements* using fragment methods and truncating the off-diagonal part of the density matrix at second nearest neighbors or third nearest neighbors. The fragments then need to overlap only a minimal amount to assemble a reasonable density matrix for the entire molecule from the fragment density matrices.

Once an approximate molecular density matrix has been built, the Coulomb field at the van der Waals surface is generated from it as in a standard ESP fitting procedure. This requires calculating the molecular Coulomb field, a step which scales asymptotically as N^2 (including the effects of integral cutoffs), where N is

proportional to the size of the molecule. However, the actually computational effort for this step is expected to be quite small (perhaps by a factor of 10-100) as compared to, e.g., an ab initio SCF (Hartree-Fock or density functional) calculation for large systems in the 100-1000-atom range. Furthermore, memory requirements are much less severe (there is no Fock matrix to store) and I/O to disk is not an issue. Consequently, we expect that our procedure will be viable even for molecules in the 1000-atom range at a reasonable computational cost.

In this paper, we do not explore these large molecule applications, focusing instead on demonstrating that the method in fact can produce accurate point charges for cases where it is possible to compare with SCF results on the entire molecule. We utilize Hartree-Fock wave functions calculated from our PSGVB[4] code for this purpose and employ the ESP fitting routines in PSGVB to determine the point charges.

Method

A. Electrostatic Potential (ESP) Method. The electrostatic potential, $V(\mathbf{r}_i)$, surrounding the molecule at a grid point[6] $\mathbf{r}_i$ is given by

$$V(\mathbf{r}_i) = \sum_{A=1}^{N_a} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} - \sum_\mu \sum_\nu P_{\mu\nu} \int \frac{x_\mu x_\nu}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}' \quad (1)$$

Here, N_a is the number of atoms in a molecule, Z_A and R_A are the charge and the position of Ath nucleus, $\mathbf{r}'$ is the position of electron, $P_{\mu\nu}$ is the density matrix element determined from the HF procedure, and x_μ and x_ν are the basis functions used. In comparison, the Mulliken population[7] is given by

$$q_A = Z_A - \sum_{\mu \in A} \sum_\nu P_{\mu\nu} S_{\mu\nu} \quad (2)$$

where q_A is the net charge on the Ath atom and $S_{\mu\nu}$ is the overlap integral. Using an atom-centered monopole approximation, the ESP, $E(\mathbf{r}_i)$ can also be expressed as

$$E(\mathbf{r}_i) = \sum_{A=1}^{N_a} \frac{q_A}{|\mathbf{r}_i - \mathbf{R}_A|} \quad (3)$$

The best least-squares fit of the net atomic charges to the ESP is obtained by minimizing the sum S

$$S(q_A) = \sum_{i=1}^{N_g} [V(\mathbf{r}_i) - E(\mathbf{r}_i)]^2 w(\mathbf{r}_i) \quad (4)$$

under the constraint

$$\sum_{A=1}^{N_a} q_A - q_{tot} = 0 \quad (5)$$