

# Derivation of Net Atomic Charges from Molecular Electrostatic Potentials

Robert J. Woods, Maged Khalil, Wendy Pell, Steven H. Moffat, and Vedene H. Smith, Jr.\*

Department of Chemistry, Queen's University, Kingston, Ontario, K7L 3N6, Canada

Received 24 March 1989; accepted 12 October, 1989

The dependence of net atomic charges, as derived from least-squares fitting to electrostatic potentials, on molecular orientation and potential site location is critically examined. A unique rotationally invariant algorithm, in which the charges may be constrained to reproduce the molecular dipole moment, is presented and the results for several basis sets compared to previous work. Significant improvements in dipole moments derived from the unconstrained charges are noted.

## INTRODUCTION

That molecular electron density may be approximated by discrete atomic electron populations has given rise to the useful concept of net atomic charge. The utility of net atomic charges in many areas of research, including Monte Carlo and molecular dynamic simulations, and molecular mechanical calculations, has led to the development of several methods for their calculation. Traditionally, these methods require that the electron density be partitioned in some manner between the constituent atoms. Well known population analysis methods include those proposed by Mulliken<sup>1</sup> and Löwdin.<sup>2</sup> Excellent reviews of the literature have been provided by Hall<sup>3</sup> and Smith.<sup>4</sup> More recently other approaches to the determination of net atomic charges have been proposed, including, direct integration of the electron density between bond critical point boundaries,<sup>5</sup> and analysis of experimental electron densities<sup>6,7</sup> and infrared spectral intensities.<sup>8</sup> A review of the use of net atomic charges in the interpretation of chemical effects has been provided by Fliszár.<sup>9</sup>

An alternate approach, currently receiving a great deal of attention, is the derivation of net atomic charges from a least-squares fitting of calculated electrostatic potentials to quantum mechanically determined electrostatic potentials.<sup>10–14</sup> This approach has the advantage that, while net atomic charge is not a rigorously defined quantity, the electrostatic potential for a point (*i*) at a position (**r**<sub>*i*</sub>) around a molecule may be explicitly defined quantum mechanically as:

$$V_i = \sum_{j=1}^{N_a} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}', \quad (1)$$

where,  $Z_j$  is the nuclear charge on atom *j* centered at **R**<sub>*j*</sub> and  $\rho$  is the electron density. Equation (1) may be rewritten in the SCF-LCAO-MO formalism as:

$$V_i = \sum_{j=1}^{N_a} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} - \sum_{\mu,\nu} P_{\mu,\nu} \int \frac{\varphi_{\mu}^*(\mathbf{r}')\varphi_{\nu}(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}', \quad (2)$$

where,  $P_{\mu,\nu}$  is the density matrix element for atomic basis functions  $\varphi_{\mu}$  and  $\varphi_{\nu}$ ,  $N_a$  is the number of atoms in the molecule, and  $N_e$  is the total number of electrons in the system. The electrostatic potential may also be expressed as an atom-centered multipole expansion. Williams has recently indicated that such an expansion, up to and including quadrupole terms, results in an excellent reproduction of the potential.<sup>14</sup> However, in order to interpret the monopole values from a multipole expansion, as simple net atomic charges, it is necessary to truncate the expansion at the monopole level. Thus, by using an atom-centered monopole approximation, the electrostatic potential at this point (*i*) may also be expressed as:

$$E_i(q_1, q_2, \dots, q_{N_a}) = \sum_{j=1}^{N_a} \frac{q_j}{|\mathbf{r}_i - \mathbf{R}_j|}, \quad (3)$$

where,  $q_j$  is the net atomic charge on atom *j*. Thus minimization of the function *Y*, where,

$$Y(q_1, q_2, \dots, q_{N_a}) = \sum_{i=1}^{N_p} w_i [V_i - E_i(q_1, q_2, \dots, q_{N_a})]^2, \quad (4)$$

\*Author to whom all correspondence should be addressed.

and where,  $N_p$  is the total number of points at which the electrostatic potential is evaluated, leads to a best least squares fit of the net atomic charges. The weighting factor,  $w_i$ , for each point is set to unity as recommended in reference 11.

An early attempt<sup>10</sup> at minimizing the function  $Y$  employed an iterative technique and required an initial guess for the charges. Cox and Williams<sup>11</sup> later solved for the minimizing charges by partial differentiation of the function  $Y$ ; for a review of their method see reference 15. A program (CHELP)<sup>16,13</sup> that uses Lagrange multipliers<sup>17</sup> to solve for the charges directly, avoiding any initial guess at their values, has been recently published. When using Lagrange multipliers, it is necessary to add constraint conditions to the function  $Y$ . This may be expressed as:

$$Z(q_1, q_2, \dots, q_{Na}, \lambda_1, \dots, \lambda_{N_c}) = Y(q_1, q_2, \dots, q_{Na}) + \sum_{k=1}^{N_c} \lambda_k G_k(q_1, q_2, \dots, q_{Na}), \quad (5)$$

where,  $N_c$  is the number of constraints, and,  $\lambda_k$  and  $G_k$  are the  $k$ th Lagrange multiplier and constraint, respectively. The most commonly imposed constraint, and the only one that is implemented in CHELP,<sup>13,16</sup> is that the sum of the net atomic charges must equal the total molecular charge, which may be expressed as,

$$G_1(q_1, q_2, \dots, q_{Na}) = \left( \sum_{j=1}^{Na} q_j \right) - Q_T = 0, \quad (6)$$

where,  $Q_T$  is the total charge on the molecule.

For our work, we have included an optional constraint, in which, the net atomic charges are forced to reproduce the quantum mechanically generated molecular dipole moment. This may be expressed in component form as:

$$G_2(q_1, q_2, \dots, q_{Na}) = \left( \sum_{j=1}^{Na} q_j C_{1j} \right) - D_1 = 0, \quad (7)$$

$$G_3(q_1, q_2, \dots, q_{Na}) = \left( \sum_{j=1}^{Na} q_j C_{2j} \right) - D_2 = 0, \quad (8)$$

$$G_4(q_1, q_2, \dots, q_{Na}) = \left( \sum_{j=1}^{Na} q_j C_{3j} \right) - D_3 = 0, \quad (9)$$

where,  $C_{mj}$  is the  $m$ th Cartesian coordinate of the  $j$ th atom and  $D_m$  is the  $m$ th component of the quantum mechanical dipole moment. Minimization of function  $Z$  by partial differentiation with respect to  $\lambda_k$  and  $q_k$  gives rise to the following expression for atom  $k$ :

$$\sum_{i=1}^{N_p} \frac{V_i}{d_{ik}} = \sum_{i=1}^{N_p} \left[ \sum_{j=1}^{Na} \frac{q_j}{d_{ij} d_{ik}} \right] + \frac{1}{2} \lambda_1 + \frac{1}{2} \lambda_2 C_{1k} + \frac{1}{2} \lambda_3 C_{2k} + \frac{1}{2} \lambda_4 C_{3k}. \quad (10)$$

These  $N_a$  eqs. (10) and the  $N_c$  constraint eqs. (6)–(9) may be written

$$\mathbb{X} = \mathbb{A}\mathbb{Y}, \quad (11)$$

where,

$$\mathbb{A} = \begin{bmatrix} \sum_i^{N_p} \frac{1}{d_{i1} d_{i1}} & \sum_i^{N_p} \frac{1}{d_{i2} d_{i1}} & \cdots & \sum_i^{N_p} \frac{1}{d_{iNa} d_{i1}} & \frac{1}{2} & \frac{1}{2} C_{11} & \frac{1}{2} C_{21} & \frac{1}{2} C_{31} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_i^{N_p} \frac{1}{d_{Np1} d_{NpNa}} & \sum_i^{N_p} \frac{1}{d_{Np2} d_{NpNa}} & \cdots & \sum_i^{N_p} \frac{1}{d_{NpNa} d_{NpNa}} & \frac{1}{2} & \frac{1}{2} C_{1Na} & \frac{1}{2} C_{2Na} & \frac{1}{2} C_{3Na} \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 & 0 \\ C_{11} & C_{12} & \cdots & C_{1Na} & 0 & 0 & 0 & 0 \\ C_{21} & C_{22} & \cdots & C_{2Na} & 0 & 0 & 0 & 0 \\ C_{31} & C_{32} & \cdots & C_{3Na} & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbb{X} = \begin{bmatrix} \sum_i^{N_p} \frac{V_i}{d_{i1}} \\ \vdots \\ \sum_i^{N_p} \frac{V_i}{d_{iNa}} \\ Q_T \\ D_1 \\ D_2 \\ D_3 \end{bmatrix}, \quad \text{and,} \quad \mathbb{Y} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_{Na} \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix}.$$

For convenience the term  $|\mathbf{r}_i - \mathbf{R}_j|$  has been represented by  $d_{ij}$ . Solution of the matrix equation,

$$\mathbf{Y} = \mathbf{A}^{-1}\mathbf{X} \quad (12)$$

then generates the net atomic charges ( $q_k$ ) and the Lagrange multipliers ( $\lambda_k$ ).

After converting CHELP to interface with the *ab initio* program GAMESS,<sup>18</sup> rather than the *ab initio* program Gaussian 82<sup>19</sup> for which it was written, we could not consistently reproduce the published<sup>13</sup> charge values (we shall refer to the version of CHELP modified for GAMESS as CHELP/GAMESS). It was subsequently determined that in order to reproduce the published<sup>13</sup> charge values the molecules had to be aligned symmetrically with respect to one or more of the cartesian axes. This requirement was necessary because CHELP evaluates the electrostatic potentials at fixed symmetric positions with respect to the coordinate system. Unlike the Gaussian series of *ab initio* programs, GAMESS does not automatically reorient the molecule symmetrically about an axis.

The use of symmetry either in point location methodology or in the evaluation of the electrostatic potentials, may be effective for the calculation of net atomic charges in small symmetric molecules. However, there are two cases when these methodologies fail. The first occurs when a symmetric molecule is not aligned symmetrically with respect to the point locations and the second occurs when the molecule is asymmetric. As our interests are in the study of asymmetric biological molecules, it was necessary to devise a new approach to the location of points, and hence, determination of net atomic charges, that had no inherent dependence on symmetry. This approach is described in the following section.

## POINT LOCATION METHODOLOGY

For each atom in the molecule, a point is chosen on the van der Waal's surface with random  $x$ -,  $y$ -, and  $z$ -coordinates with respect to the coordinate origin. This point is then extended randomly to between the van der Waal's radius  $+a$  Å and the van der Waal's radius  $+b$  Å (for simplicity we shall refer to this shell as van der Waal's  $[a, b]$ ). A second point is generated by the inversion of this point through the origin. This gives a pair of points  $(x, y, z)$  and  $(-x, -y, -z)$ , which has no net displacement about the origin. A total of  $N_p$  points is thus generated for each atom in the structure. Each group of  $N_p$  points is then translated to the relevant nuclear position and any points which lie within the van der Waal's radius of any neighbouring atoms are discarded.

The array of points may then be used as input to any computer program capable of generating electrostatic potentials at these locations around the molecule. We have made use of the program GAMESS<sup>19</sup> for the SCF evaluation of all the quantum mechanical quantities, although there are several other packages capable of performing these calculations.

We have chosen to store the points in a nonbinary format and not to incorporate the electrostatic potential calculation within the program so as to ensure program flexibility, while at the same time minimizing system dependence. The program, PDQP, (Potentially Derived Quantities Point program) is currently written to store the points in the correct format for input into the program GAMESS. However, PDQP may easily be modified to store the points in any other format.

Once the electrostatic potentials have been evaluated, they are used as input for the charge fitting program, PDQC (Potentially Derived Quantities Charge program). PDQC carries out least squares fitting of the electrostatic potentials, using Lagrange multipliers to constrain the atomic charges such that they reproduce the total molecular charge and, optionally, the molecular dipole moment.

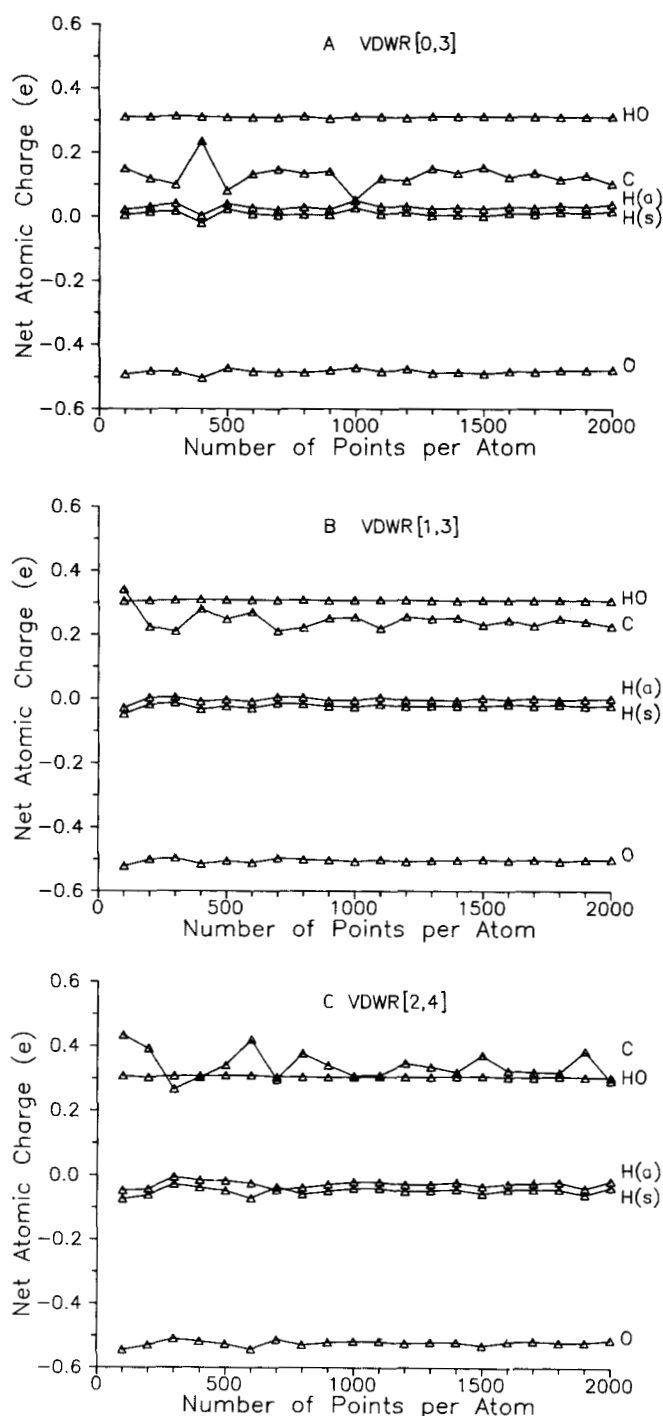
## RESULTS AND DISCUSSION

The following section includes discussions of several factors affecting the charge calculations. The dependence of the calculations on the number of points at which the electrostatic potentials are evaluated, and on the molecular orientation is examined. Results for electrostatic potentials calculated with several basis sets for a variety of molecules and anions are included. All calculations were performed without the inclusion of dipole constraints, unless otherwise stated.

### Number of Points

Figure 1 illustrates the dependence of net atomic charge on the number of points used in the least squares fitting procedure. The electrostatic potentials for the experimental geometry of methanol were calculated with an STO-3G<sup>20</sup> basis set. For each data set a different random number 'seed'<sup>21</sup> was used in the random number generator<sup>22</sup> to ensure a unique array of points for each calculation. Although no symmetry restrictions have been imposed, the charges on the staggered hydrogen atoms are virtually indistinguishable and the charge on only one of them is plotted in Figure 1.

The net atomic charges begin to stabilize beyond 1000 points per atom for all the atoms.



**Figure 1.** Dependence of net atomic charge for the atoms in methanol on the number of points and the point locations. (a) all points located between van der Waal's [0,3]; (b) all points located between van der Waal's [1,3]; (c) all points located between van der Waal's [2,4]. Electrostatic potentials were calculated using the STO-3G basis set. Labels refer to the atoms as follows: C, carbon; HO, hydroxyl hydrogen; H(a), anti-hydrogen; H(s), staggered hydrogen; and O, oxygen.

However, the carbon atom shows a greater fluctuation under 1000 points per atom than any of the other atoms. This atom has more substituents than the other atoms and so it will have a smaller fraction of its van der Waal's surface ex-

posed. This will result in there being fewer points located close to the carbon atom and, consequently, the representation of the electrostatic potential near it will be poor. Thus, more points are required in order to calculate a well defined

net atomic charge for this atom. Furthermore, the increased variation in the carbon atomic charge may be viewed as being a response to the smaller fluctuations in each of the hydrogen atomic charges.

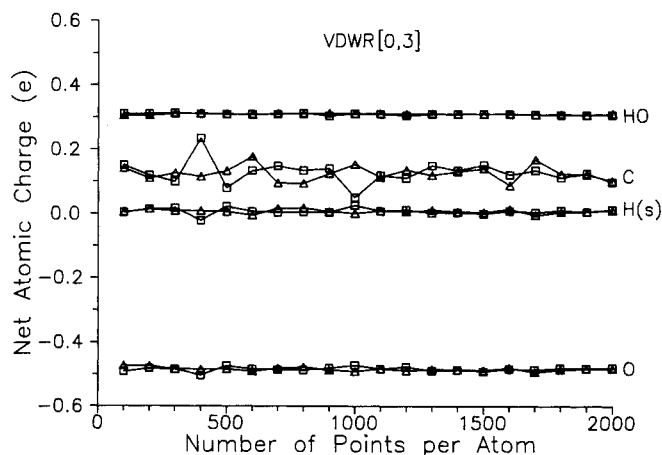
In order to examine the validity of the monopole approximation, within the framework of our methodology, electrostatic potentials were evaluated not only between van der Waal's [0, 3] as shown in Figure 1(a), but also, at distances farther from the molecule. The results for electrostatic potentials evaluated between van der Waal's [1, 3] and van der Waal's [2, 4] are illustrated in Figures 1(b) and (c), respectively. At 2000 points per atom, the net atomic charges on the staggered hydrogen atoms decrease from 0.016e at van der Waal's [0, 3] to  $-0.021e$  at van der Waal's [1, 3] to  $-0.036e$  at van der Waal's [2, 4] and similarly the charge on the anti-hydrogen atom decreases from 0.031e to 0.001e to  $-0.016e$ .

As the electrostatic potentials are evaluated farther from the van der Waal's surface of the carbon atom, there are smaller positive contributions from the carbon nucleus. Thus, the net atomic charges on the hydrogen atoms decrease artificially as the points are located farther from the van der Waal's surface. As the hydrogen atomic charges decrease, the carbon atomic charge increases from 0.101e at van der Waal's [0, 3] to 0.227e at van der Waal's [1, 3] to 0.295e at van der Waal's [2, 4] at 2000 points per atom, thus, conserving total molecular charge.

In order to insure a reasonable representation of the charges on all atoms, electrostatic potentials evaluated near the van der Waal's surface, such as between van der Waal's [0, 3], need to be included in the least squares fitting routine. This is especially important for those atoms that are less exposed, such as the carbon atom in methanol. This generalization has been examined for the cases of four moderately large molecules containing differing functionalities, *vide infra*.

### Rotational Dependence

Figure 2 addresses the question of rotational variance as applied to the charge on the atoms in methanol with the electrostatic potentials evaluated between van der Waal's [0, 3]. The graph shows the charges calculated for the atoms in symmetrically aligned methanol and in methanol that was rotated arbitrarily  $10^\circ$  about each of the Cartesian axes. For clarity, the anti-hydrogen atom is omitted from the figure. That there is no dependence on molecular orientation is supported by the observation that there is no net offset between the limiting charge values for any of the atoms in either orientation; both sets of charges fluctuate about the same values and de-



**Figure 2.** Dependence of net atomic charge for the atoms in methanol on the number of points for symmetric ( $\Delta$ ) and rotated ( $\square$ ) orientations of methanol. Electrostatic potentials were calculated using the STO-3G basis set at points located between van der Waal's [0, 3]. Labels refer to the atoms as follows: C, carbon; HO, hydroxyl hydrogen; H( $\alpha$ ), anti-hydrogen; H(s), staggered hydrogen; and O, oxygen.

pend on the total number of points in analogous manners. A change in molecular orientation affects the charges in a manner similar to a change in the random number seed. The charge on the carbon atom in the symmetrically aligned molecule at 2000 points per atom is 0.101e and the corresponding value for the rotated geometry is 0.104e.

Table I compares the values for the charges on the atoms in water and methanol for both symmetric and rotated orientations as calculated by PDQC at 2000 points per atom and CHELP/GAMESS at 250 points per molecule as recommended in reference 13. Increasing the number of points used in CHELP/GAMESS from 250 to 1000 does not cause convergence of the charge values in the symmetric and rotated orientations of the molecules. Consequently, only the results for CHELP/GAMESS, in which, 250 points are used are reported. Also listed are the charge values reported in reference 11. PDQC consistently exhibits a smaller difference between the values of the charges in the symmetric and rotated orientations than CHELP/GAMESS; however, neither method shows significant differences in the atomic charges calculated for the symmetric and rotated orientations of water. The absolute values for these charges are in good agreement with each reference value shown in Table I.

There appears to be a more significant difference between the charges calculated for the symmetric and rotated orientations of methanol than for water. Notably, the absolute value of the charge on the carbon atom, in either orientation, as calculated by PDQC is significantly lower (by

**Table I.** Net atomic charges at STO-3G basis level using experimental geometries.

Atom	PDQC <sup>a</sup>		CHELP/GAMESS <sup>c</sup>		Reference 11
H <sub>2</sub> O					
O	-0.613	-0.613 <sup>b</sup> (0.000) <sup>d</sup>	-0.619	-0.618 <sup>b</sup> (0.001)	-0.614
H	0.307	0.306(0.001)	0.310	0.309(0.001)	0.307
H	0.306	0.307(0.001)	0.310	0.309(0.001)	0.307
rms <sup>e</sup>	0.309	0.232	0.142	0.158	0.120
CH <sub>3</sub> OH					
C	0.101	0.104(0.003)	0.192	0.198(0.006)	0.188
O	-0.479	-0.482(0.003)	-0.490	-0.510(0.020)	-0.501
HO	0.310	0.313(0.003)	0.312	0.320(0.008)	0.317
H(a)	0.037	0.039(0.002)	0.012	0.018(0.006)	0.014
H(s)	0.016	0.013(0.003)	-0.013	-0.013(0.000)	-0.009
H(s)	0.015	0.013(0.002)	-0.013	-0.014(0.001)	-0.009
CH <sub>3</sub>	0.169	0.169(0.000)	0.178	0.189(0.011)	0.184
rms <sup>e</sup>	0.754	0.753	0.985	0.818	0.382

<sup>a</sup> Calculations performed at 2000 points per atom located between van der Waal's [0, 3].

<sup>b</sup> Geometry rotated 10° about each axis.

<sup>c</sup> Calculations performed at 250 points per molecule.

<sup>d</sup> Values in parentheses represent the difference between charges calculated for symmetric and rotated structures.

<sup>e</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

about 0.090e) than either equivalent value as calculated by CHELP/GAMESS or the values from reference 11. The staggered hydrogen atoms carry partial positive charges, as calculated by PDQC, while the other methods generate partial negative charges of about the same absolute magnitude. Furthermore, PDQC predicts that the anti-hydrogen atom carries a larger partial positive charge than predicted by either of the other methods.

As already discussed, the location of the points about the molecule plays a key role in determining the value of the charges on these hydrogen atoms. In order for valid comparisons to be made between charges calculated by different methodologies, it is essential that the electrostatic potentials are evaluated within the same spatial boundaries around each molecule. Consequently, it is useful to examine the total charge on the methyl group when comparing results from several sources. If this is done, it becomes apparent that there is little difference between the rotated and symmetric orientations as calculated by PDQC while there appears to be a notable difference between the analogous values calculated by CHELP/GAMESS. However, the difference between the values calculated by PDQC and those calculated by CHELP/GAMESS is now much smaller (approximately 0.015e) than observed for the carbon atoms alone.

### Basis Sets

The effects of employing a larger basis set (6-31G\*\*<sup>23</sup>) in the evaluation of the electrostatic

potentials is illustrated in Table II. As in Table I, both water and methanol in symmetric and rotated orientations are used for charge comparisons. The charges calculated by CHELP/GAMESS as well as those given in references 11 and 12 have been included for comparison. As seen for the STO-3G results, the values for the charges on the atoms in water are grossly consistent regardless of methodologies although the average error between the values for the charges in the symmetric and rotated orientations is slightly larger than found in Table I. For methanol at 6-31G\*\*, PDQC again yields a smaller charge on the carbon atom than do the other methods and at 2000 points per atom there is less difference between the charges on the atoms in the symmetric and rotated orientations in our results. Analogous to the STO-3G results, there is less difference in the total charges on the methyl groups between the two orientations for both our results and CHELP/GAMESS values than for the carbon atomic charges alone. Also included in Table II are the results for PDQC calculations at 2500 points per atom.

With current methods,<sup>10-14</sup> increasing the number of points at which the electrostatic potentials are evaluated may result in those points being located too far away from the molecule to contribute significantly to the potential. This creates an artificial convergence of the charge values and a noticeable, but, uninformative lowering of the rms deviation.<sup>24</sup> With PDQC all potentials are evaluated within a fixed boundary around the molecule. Consequently, an increase

**Table II.** Net atomic charges at 6-31G\*\* basis level using experimental geometries.

Atom	PDQC <sup>a</sup>		CHELP/GAMESS <sup>c</sup>		Reference 11	Reference 12
H <sub>2</sub> O						
O	-0.788	-0.785 <sup>b</sup> (0.003) <sup>d</sup>	-0.798	-0.792 <sup>b</sup> (0.006)	-0.786	-0.794
H	0.395	0.391 (0.004)	0.399	0.395 (0.004)	0.393	0.397
H	0.395	0.394 (0.001)	0.399	0.396 (0.001)	0.393	0.397
rms <sup>e</sup>	0.695	0.694	0.648	0.648	0.502	1.540
CH <sub>3</sub> OH						
C	0.182	0.148 (0.038)	0.236	0.302 (0.066)	0.244	0.149
O	-0.657	-0.652 (0.005)	-0.652	-0.689 (0.037)	-0.680	-0.656
HO	0.415	0.416 (0.001)	0.407	0.419 (0.012)	0.424	0.424
H(a)	0.061	0.070 (0.009)	0.042	0.036 (0.006)	0.048	-0.001
H(s)	0.000	0.008 (0.008)	-0.017	-0.034 (0.018)	-0.018	0.042
H(s)	-0.001	0.010 (0.011)	-0.017	-0.035 (0.018)	-0.018	0.042
CH <sub>3</sub>	0.242	0.236 (0.006)	0.244	0.269 (0.025)	0.256	0.232
rms <sup>e</sup>	0.860	0.860	1.003	0.853	0.382	1.400
CH <sub>3</sub> OH <sup>f</sup>						
C	0.175	0.174 (0.001)				
O	-0.684	-0.682 (0.002)				
HO	0.441	0.439 (0.002)				
H(a)	0.070	0.074 (0.004)				
H(s)	-0.001	-0.003 (0.002)				
H(s)	-0.002	-0.003 (0.001)				
CH <sub>3</sub>	0.242	0.242 (0.000)				
rms <sup>e</sup>	1.201	1.201				

<sup>a</sup> Calculations performed at 2000 points per atom located between van der Waal's [0, 3].

<sup>b</sup> Geometry rotated 10° about each axis.

<sup>c</sup> Calculations performed at 250 points per molecule.

<sup>d</sup> Values in parentheses represent the difference between charges calculated for symmetric and rotated structures.

<sup>e</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

<sup>f</sup> Calculations performed at 2500 points per atom located between van der Waal's [0, 3].

in the number of points may have a significant effect on the value of the charges and may lead to a lowering of the root mean square value directly related to the goodness of fit. In the case of methanol, an increase from 2000 points per atom to 2500 points per atom, greatly reduces the dif-

ferences between the values of the atomic charges calculated for the symmetric and rotated molecular orientations.

A comparison of results for anions is presented in Table III. The PDQC calculations were performed at both 1000 and 2000 points per atom

**Table III.** Net atomic charges on anions<sup>a</sup> calculated at various basis sets.

Atom	3-21G		6-31 + G		6-31G**	
	2000 <sup>b</sup>	1000 <sup>c</sup>	2000	1000	2000	1000
OH <sup>-</sup>						
O	-1.134	-1.135 (-1.095) <sup>d</sup>	-1.327	-1.347 (-1.305)	-1.170	-1.170 (-1.171)
H	0.134	0.135 ( 0.096)	0.327	0.347 ( 0.305)	0.170	0.170 ( 0.171)
rms <sup>e</sup>	0.441	0.447 ( 0.376)	1.450	1.474 ( 1.930)	0.757	0.588 ( 1.258)
CN <sup>-</sup>						
C	-0.539	-0.545 (-0.537)	-0.506	-0.506 (-0.489)	-0.509	-0.515 (-0.504)
N	-0.461	-0.455 (-0.463)	-0.494	-0.494 (-0.511)	-0.491	-0.487 (-0.496)
rms <sup>e</sup>	1.892	1.783 ( 2.300)	2.081	2.045 ( 2.797)	1.640	1.560 ( 1.706)

<sup>a</sup> Geometries optimized at each basis set.

<sup>b</sup> 2000 points per atom located between van der Waal's [0, 3].

<sup>c</sup> 1000 points per atom located between van der Waal's [0, 3].

<sup>d</sup> Values in parentheses are from reference 13 (CHELP at 250 points). The RMS values given in reference 13 are incorrect and the corrected values are given above. See also reference 24.

<sup>e</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

located between van der Waal's [0,3]. For the hydroxide ion our values and those from reference 13 are in reasonable agreement, although, for the calculations performed at the 3-21G<sup>25</sup> and 6-31 + G<sup>26</sup> levels, PDQC predicts slightly larger absolute values for the atomic charges. In the case of the cyanide ion PDQC predicts the carbon atom to be more negative than the nitrogen atom, for each basis set. Interestingly, reference 13 gives values for the cyanide ion at the 6-31 + G level in which the atomic charges appear reversed; the nitrogen atom is the more negative atom. Our calculations using CHELP/GAMES, verified these results and we conclude that this reversal is indicative of the need for a higher density of points than is possible when using CHELP. The point location methodology in CHELP is such that all points are located along only 14 vectors, originating from each atom and directed along either the Cartesian axes or along the vector passing through the center of each octant. The points are placed in shells 1 Å apart beginning at the van der Waal's surface and there is no allowance for increasing the number of vectors. Such a point distribution is characterized by constant linear density along any of the vectors, but, has a spatial point density that decreases as a function of the square of the distance of the point from the nucleus. Previously,<sup>13</sup> it has been noted that the 6-31 + G basis set give rises to larger rms deviations than either the 3-21G or 6-31G\*\* basis sets, when evaluated at the same number of points. Although this is also apparent for PDQC, the increase in the rms values for the 6-31 + G results is smaller. Accordingly, the PDQC charges on the two atoms in the cyanide ion maintain the same relationship, with the carbon atom being the more negative, regardless of the basis set used in the calculation of the electrostatic potentials.

## Dipoles

Molecular dipole moments calculated for water and methanol with STO-3G and 6-31G\*\* basis sets are shown in Table IV. The results indicate

**Table IV.** Dipole moments<sup>a</sup> (D).

Molecule (basis set)	PDQC <sup>b</sup>	SCF <sup>c</sup>	Reference 13	Reference 12	Reference 11
H <sub>2</sub> O (STO-3G)	1.726	1.725	1.746	1.732	1.727
H <sub>2</sub> O (6-31G**)	2.220	2.186	2.258	2.242	2.211
CH <sub>3</sub> OH (STO-3G)	1.500	1.492	1.440	—	1.500 <sup>d</sup>
CH <sub>3</sub> OH (6-31G**)	1.903	1.914	1.864	2.013	1.899

<sup>a</sup> Calculations performed at experimental geometries.

<sup>b</sup> Calculations performed at 2000 points per atom located between van der Waal's [0, 3].

<sup>c</sup> Quantum mechanical (SCF) calculations were performed using the program GAMESS.

<sup>d</sup> This value appears to be transposed with the SCF-derived dipole moment in reference 11.

that PDQC can lead to extremely good reproduction of the quantum mechanical dipole moments. Obviously, the accuracy of the dipole moment reproduction is controlled by the number of points used in the calculations.

The effects of constraining the atomic charges to reproduce the quantum mechanical dipole moment is illustrated in Table V for methanol for electrostatic potentials evaluated between van der Waal's [0,3] with both STO-3G and 6-31G\*\* basis sets. The constrained charges differ only slightly from the values calculated without the dipole constraints as shown in Tables I and II. In fact, the dipole moments calculated without constraints are already so close to the quantum mechanical values that constraining the charges has no effect on the root mean square values. Consequently, the most significant effect of the dipole constraints is to remove any asymmetry among the charges in symmetric molecules.

## Larger Molecules

The results for charge calculations on two monosaccharides,  $\beta$ -D-fructopyranose<sup>27</sup>(1) and its ring

**Table V.** PDQC<sup>a</sup> net atomic charges calculated with dipole constraints.

Atom	STO-3G		6-31G**	
	Symmetric	Rotated <sup>b</sup>	Symmetric	Rotated <sup>b</sup>
CH <sub>3</sub> OH				
C	0.099	0.104	0.183	0.149
O	-0.477	-0.482	-0.658	-0.652
HO	0.311	0.315	0.418	0.419
H(a)	0.036	0.037	0.058	0.067
H(s)	0.015	0.013	-0.001	0.009
H(s)	0.015	0.013	-0.001	0.009
rms <sup>c</sup>	0.754	0.753	0.860	0.860

<sup>a</sup> Calculations performed at 2000 points per atom located between van der Waal's [0, 3].

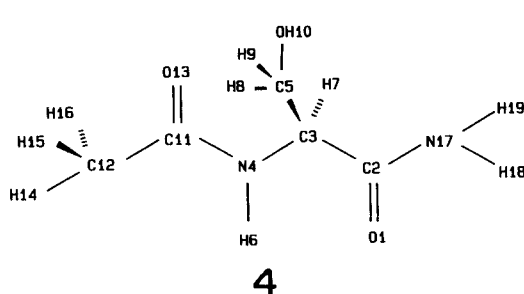
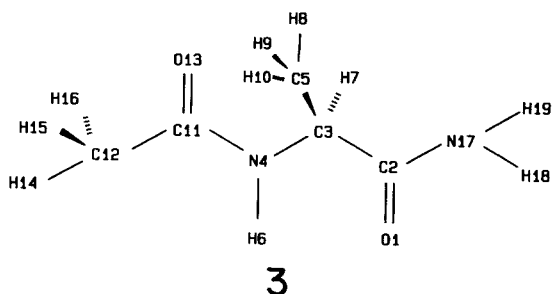
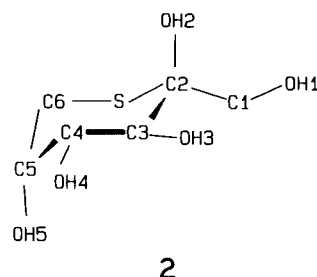
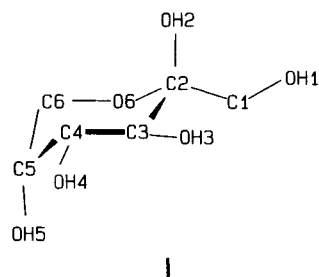
<sup>b</sup> Geometry rotated 10° about each axis.

<sup>c</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.



sulphur analogue, 6-thio- $\beta$ -D-fructopyranose<sup>28</sup> (2), as well as for two amino acids, *N*-acetyl-L-alanyl amide<sup>29</sup> (3) and *N*-acetyl-L-serinyl amide<sup>30</sup> (4) are presented in Tables VI–IX, respectively. Each of these natural products is asymmetric. The charge calculations have been carried out at varying shell thicknesses and point densities. Included for comparison are the Mulliken charges<sup>31</sup> as well as charges obtained using CHELP/GAMESS. To insure that the CHELP/GAMESS charges reached their limiting values, several calculations were carried out for each structure with progressively increasing numbers of points.

On examination of the atomic charges in 1, calculated by PDQC for points located between van der Waal's [0, 3], it becomes evident that the atoms may be classified into groups in terms of charge. The hydroxyl oxygen atoms are the most negative followed by the ring oxygen atom. The non-hydroxyl hydrogen atoms are approximately neutral while the hydroxyl hydrogen atoms are the most positive. The carbon atoms vary significantly in charge. Not surprisingly, the most positive carbon atom is C-2, which is bonded to both O-2 and O-6. The tertiary carbon atoms are less positive, while primary carbon atoms C-1 is nearly neutral. Interestingly, the other primary carbon atom, C-6, carries a partial negative charge.



The Mulliken charges illustrate similar trends, however, the absolute values are in general smaller than the PDQC values. Furthermore, a close examination of the Mulliken charges reveals that although C-2 is still the most positive carbon atom, the relative charges on the remaining carbon atoms do not parallel those found by PDQC. Similarly, while O-2 is the most negative oxygen atom, the Mulliken charges on the other

hydroxyl oxygen atoms do not follow the same ordering as predicted by PDQC.

The more exposed atoms, such as the hydroxyl oxygen and hydrogen atoms, show the least differences between the PDQC and CHELP/GAMESS charges. There are, however, notable differences between our results and the CHELP/GAMESS results for the carbon and non-hydroxyl hydrogen atoms. While CHELP/GAMESS does predict C-2 to be the most positive carbon atom and C-6 to be negative, the CHELP/GAMESS values for the charges on atoms C-1, C-3, and C-5 do not agree with PDQC. These charges differ by approximately  $\pm 0.5e$  from the PDQC values, with the biggest difference being found for the charge on C-3.

That the differences between the PDQC and CHELP/GAMESS charges are related to point location may be illustrated by the PDQC results for charges derived from electrostatic potentials evaluated between van der Waal's [1, 3]. By omitting points located within 1 Å of the van der Waal's surface of the molecule the most significant effects on the charges are found for the carbon and non-hydroxyl hydrogen atoms. These results predict an unlikely ordering of the carbon atomic charges, in which, C-2 is no longer the most positive carbon atom and C-6 carries more

negative charge than the ring oxygen atom, O-6. The relative ordering of the charges on the exposed hydroxyl oxygen and hydrogen atoms appears to be much less dependent on the inclusion or exclusion of electrostatic potentials evaluated close to the molecule.

Table VII contains the PDQC, CHELP/GAMESS, and Mulliken charges for 2. A similar classification of the atoms in terms of charge may be made

**Table VI.**  $\beta$ -D-Fructopyranose<sup>a</sup> net atomic charges at STO-3G basis level.

Atom	CHELP/GAMESS			2000 <sup>c</sup>	PDQC			Mulliken <sup>e</sup>
	500 <sup>b</sup>	1000 <sup>b</sup>	2000 <sup>b</sup>		1500 <sup>c</sup>	1000 <sup>c</sup>	1000 <sup>d</sup>	
C <sub>6</sub> H <sub>12</sub> O <sub>6</sub>								
C-1	0.051	0.070	0.075	0.013	0.010	0.043	-0.041	-0.005
C-2	0.523	0.515	0.511	0.471	0.459	0.453	0.279	0.239
C-3	0.287	0.290	0.294	0.351	0.345	0.342	0.346	0.060
C-4	0.392	0.390	0.389	0.332	0.320	0.309	0.130	0.066
C-5	0.099	0.098	0.097	0.192	0.212	0.207	0.221	0.066
C-6	-0.098	-0.101	-0.110	-0.078	-0.132	-0.101	-0.386	0.003
O-1	-0.458	-0.465	-0.467	-0.486	-0.483	-0.486	-0.459	-0.317
O-2	-0.595	-0.592	-0.590	-0.594	-0.595	-0.596	-0.538	-0.331
O-3	-0.580	-0.581	-0.582	-0.591	-0.580	-0.578	-0.539	-0.313
O-4	-0.577	-0.575	-0.574	-0.577	-0.574	-0.566	-0.524	-0.322
O-5	-0.489	-0.488	-0.487	-0.518	-0.515	-0.514	-0.479	-0.318
O-6	-0.320	-0.321	-0.317	-0.344	-0.322	-0.327	-0.178	-0.285
H-1	0.026	0.020	0.019	0.049	0.055	0.045	0.083	0.070
H-1'	0.027	0.024	0.023	0.047	0.045	0.035	0.072	0.055
H-3	-0.015	-0.016	-0.017	-0.024	-0.019	-0.021	0.012	0.073
H-4	-0.036	-0.036	-0.036	-0.025	-0.018	-0.012	0.054	0.064
H-5	0.026	0.026	0.029	-0.002	-0.003	-0.004	0.039	0.052
H-6	0.071	0.073	0.075	0.061	0.075	0.066	0.153	0.064
H-6'	0.098	0.099	0.101	0.095	0.108	0.099	0.167	0.080
HO-1	0.272	0.275	0.275	0.307	0.305	0.302	0.301	0.209
HO-2	0.324	0.322	0.321	0.333	0.335	0.336	0.333	0.196
HO-3	0.327	0.328	0.328	0.331	0.324	0.325	0.319	0.195
HO-4	0.326	0.326	0.325	0.328	0.326	0.323	0.323	0.200
HO-5	0.319	0.319	0.319	0.325	0.321	0.321	0.315	0.210
Dipole <sup>f</sup>	2.116	2.113	2.114	2.124	2.124	2.119	2.104	
Rms <sup>g</sup>	1.006	1.012	0.805	0.991	0.859	0.981	1.579	

<sup>a</sup> Geometry from reference 27.<sup>b</sup> Total number of points per molecule.<sup>c</sup> Number of points per atom located between van der Waal's [0, 3].<sup>d</sup> Number of points per atom located between van der Waal's [1, 3].<sup>e</sup> Values taken from reference 31.<sup>f</sup> SCF - dipole = 2.105438 D.<sup>g</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

as was observed for **1**. In the case of **2**, the substitution of a sulphur atom for the ring oxygen atom in **1** affects the ordering of both the carbon and oxygen charges. The most positive carbon atom is no longer C-2, but, rather, C-4. Furthermore, the most negative oxygen atom is now O-4 and not O-2.

The charges calculated for **2** using CHELP/GAMESS again differ significantly both for the carbon atoms and for the non-hydroxyl hydrogen atoms when compared to the PDQC results. The carbon charges differ from the PDQC values by approximately  $\pm 0.5e$  with the largest differences associated with the charges on atoms C-1 and C-4.

The charges on the atoms in **3** are presented in TABLE VIII. The CHELP/GAMESS charges reach their limiting values at about 600 points per molecule. Most of the values are very similar to those values determined by PDQC. However, the charge on atom C-5 is more negative, by ap-

proximately  $0.2e$ , as calculated by PDQC. Accordingly, the PDQC charges on each hydrogen atom attached to atom C-5 are more positive by about  $0.07e$ . The differences are somewhat greater when the electrostatic potentials are evaluated between van der Waal's [1, 3].

Unlike the atomic charges in **3** calculated by PDQC, the atomic charges in **4** do not exhibit the same general trends as predicted by the other methodologies, see Table IX. While the PDQC charges calculated for points located between van der Waal's [0, 3] exhibit similar overall trends as predicted by CHELP/GAMESS, the same is not the case for PDQC charges calculated for points located between van der Waal's [1, 3]. In particular the charges on atoms C-2, C-3, N-4, and C-5 calculated by PDQC for points located between van der Waal's [1, 3] vary significantly from the other analogous values. The  $0.25e$  increase in charge on C-3 is accompanied by a drop in the charges on the neighbouring atoms, C-2

**Table VII.** 6-Thio- $\beta$ -D-Fructopyranose<sup>a</sup> atomic charges at STO-3G basis level.

Atom	CHELP/GAMESS			2000 <sup>c</sup>	PDQC			Mulliken <sup>e</sup>
	500 <sup>b</sup>	1000 <sup>b</sup>	2000 <sup>b</sup>		1500 <sup>c</sup>	1000 <sup>c</sup>	1000 <sup>d</sup>	
C <sub>6</sub> H <sub>12</sub> O <sub>5</sub> S								
C-1	0.300	0.296	0.290	0.180	0.135	0.118	0.022	-0.021
C-2	0.367	0.362	0.367	0.302	0.393	0.360	0.180	0.044
C-3	0.069	0.096	0.090	0.168	0.100	0.090	0.032	0.053
C-4	0.581	0.541	0.540	0.446	0.505	0.479	0.380	0.053
C-5	0.021	0.028	0.013	0.092	0.039	0.090	0.123	0.068
C-6	-0.026	-0.027	-0.027	-0.010	0.085	-0.006	-0.111	-0.212
O-1	-0.524	-0.530	-0.528	-0.502	-0.505	-0.493	-0.464	-0.306
O-2	-0.470	-0.478	-0.476	-0.487	-0.503	-0.488	-0.418	-0.317
O-3	-0.555	-0.554	-0.553	-0.546	-0.545	-0.529	-0.483	-0.315
O-4	-0.601	-0.587	-0.584	-0.566	-0.574	-0.573	-0.540	-0.309
O-5	-0.443	-0.450	-0.448	-0.464	-0.457	-0.463	-0.463	-0.308
S	-0.217	-0.224	-0.225	-0.201	-0.249	-0.216	-0.143	0.098
H-1	-0.054	-0.051	-0.049	-0.018	-0.002	0.005	0.047	0.059
H-1'	-0.025	-0.021	-0.020	0.028	0.035	0.038	0.090	0.077
H-3	0.006	0.001	-0.001	-0.010	-0.003	0.008	0.064	0.049
H-4	-0.093	-0.086	-0.084	-0.060	-0.073	-0.068	-0.027	0.070
H-5	0.052	0.052	0.057	0.034	0.035	0.029	0.041	0.079
H-6	0.090	0.087	0.089	0.069	0.054	0.72	0.101	0.075
H-6'	0.004	0.015	0.016	0.011	-0.017	0.012	0.043	0.076
HO-1	0.318	0.323	0.322	0.317	0.323	0.318	0.320	0.200
HO-2	0.291	0.297	0.297	0.317	0.318	0.314	0.313	0.212
HO-3	0.328	0.324	0.325	0.318	0.321	0.315	0.305	0.195
HO-4	0.327	0.326	0.325	0.318	0.320	0.323	0.318	0.191
HO-5	0.252	0.261	0.265	0.264	0.264	0.264	0.270	0.177
Dipole <sup>f</sup>	3.905	3.894	3.901	3.910	3.911	3.907	3.904	
Rms <sup>g</sup>	1.207	1.296	0.939	1.570	1.202	1.402	2.583	

<sup>a</sup> Geometry from reference 28.<sup>b</sup> Total number of points per molecule.<sup>c</sup> Number of points per atom located between van der Waal's [0, 3].<sup>d</sup> Number of points per atom located between van der Waal's [1, 3].<sup>e</sup> Values taken from reference 31.<sup>f</sup> SCF - dipole = 3.905367 D.<sup>g</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

and N-4, of about 0.1e each. The more positive charges on atoms C-2 and N-4 appear to neutralize the more negative charge on atom C-3 and the less positive charge on atom C-5. These results are illustrative of the change in the relative ordering of the atomic charges which is a consequence of neglecting the evaluation of electrostatic potentials near the molecule.

The most positively charged atoms amongst all three methodologies, i.e., PDQC, CHELP/GAMESS, and Mulliken, in both modified amino acid molecules, are C-2 and C-11, the carbonyl carbon atoms. The acetyl carbonyl carbon atom, C-11, is consistently more positive than the amide carbonyl carbon atom, C-2, except for the PDQC calculation of 4 in which the electrostatic potential is evaluated between van der Waal's [1, 3]. Notably, the most negatively charged atoms in both molecules are N-4 and N-17. The most negatively charged oxygen atom in 4 is O-6, the hydroxyl oxygen atom.

## CONCLUSIONS

To insure rotational invariance in charge calculations based on the algorithms described in this paper, it is necessary to evaluate the electrostatic potentials at between 1500 to 2500 points per atom. The inclusion of points between van der Waal's [0, 1] is essential for chemically reasonable charges and the location of points between van der Waal's [0, 3] gives results most consistent with current literature values. When working with symmetric molecules the use of dipole constraints is recommended.

We thank the Natural Sciences and Engineering Research Council of Canada for financial support in the form of a grant (to V. H. S) and a scholarship (to R. J. W.) and the Province of Ontario Ministry of Colleges and Universities for the award of scholarships (to M. K. and W. P.). The authors wish to thank L. E. Chirlian and M. M. Francl for their generous assistance in clarifying technical details pertaining to CHELP.

**Table VIII.** N-acetyl-L-alanyl amide<sup>a</sup> atomic charges at STO-3G basis level.

Atom	CHELP/GAMESS			PDQC		Mulliken <sup>e</sup>
	300 <sup>b</sup>	600 <sup>b</sup>	900 <sup>b</sup>	1500 <sup>c</sup>	1500 <sup>d</sup>	
C <sub>5</sub> H <sub>10</sub> N <sub>2</sub> O <sub>2</sub>						
O-1	-0.464	-0.474	-0.475	-0.464	-0.475	-0.297
C-2	0.633	0.694	0.698	0.683	0.720	0.285
C-3	0.265	0.182	0.177	0.203	0.198	0.053
N-4	-0.642	-0.631	-0.628	-0.677	-0.649	-0.357
C-5	-0.310	-0.248	-0.234	-0.439	-0.507	-0.170
H-6	0.281	0.279	0.277	0.296	0.281	0.189
H-7	0.012	0.033	0.033	0.051	0.056	0.100
H-8	0.097	0.086	0.081	0.134	0.148	0.071
H-9	0.067	0.053	0.050	0.108	0.127	0.066
H-10	0.101	0.085	0.082	0.139	0.157	0.073
C-11	0.739	0.741	0.739	0.816	0.769	0.298
C-12	-0.485	-0.447	-0.445	-0.589	-0.416	-0.197
O-13	-0.477	-0.485	-0.485	-0.500	-0.507	-0.314
H-14	0.113	0.103	0.102	0.138	0.094	0.072
H-15	0.109	0.100	0.100	0.139	0.096	0.063
H-16	0.125	0.112	0.112	0.145	0.102	0.081
N-17	-0.782	-0.820	-0.820	-0.828	-0.860	-0.422
H-18	0.347	0.350	0.350	0.355	0.364	0.216
H-19	0.271	0.287	0.287	0.292	0.301	0.190
Dipole <sup>f</sup>	3.531	3.554	3.548	3.518	3.493	
Rms <sup>g</sup>	0.970	1.029	0.960	0.802	0.167	

<sup>a</sup> Geometry from reference 29.<sup>b</sup> Total number of points per molecule.<sup>c</sup> Number of points per atom located between van der Waal's [0, 3].<sup>d</sup> Number of points per atom located between van der Waal's [1, 3].<sup>e</sup> Values calculated using the program GAMESS.<sup>f</sup> SCF - dipole = 3.492775 D.<sup>g</sup> Root mean square deviation between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

**Table IX.** N-acetyl-L-serinyl amide<sup>a</sup> atomic charges at STO-3G basis level.

Atom	CHELP/GAMESS			PDQC		Mulliken <sup>e</sup>
	300 <sup>b</sup>	600 <sup>b</sup>	900 <sup>b</sup>	1500 <sup>c</sup>	1500 <sup>d</sup>	
C <sub>5</sub> H <sub>10</sub> N <sub>2</sub> O <sub>2</sub>						
O-1	-0.479	-0.479	-0.478	-0.455	-0.466	-0.286
C-2	0.784	0.795	0.791	0.728	0.817	0.287
C-3	0.060	0.020	0.010	0.045	-0.212	0.053
N-4	-0.699	-0.685	-0.681	-0.655	-0.534	-0.356
C-5	0.161	0.207	0.226	0.155	0.078	0.013
H-6	0.317	0.315	0.315	0.305	0.276	0.188
H-7	0.049	0.056	0.058	0.050	0.122	0.091
H-8	0.033	0.020	0.013	0.030	0.069	0.058
H-9	0.010	0.004	0.001	0.022	0.062	0.066
O-10	-0.525	-0.531	-0.533	-0.525	-0.499	-0.308
C-11	0.821	0.819	0.817	0.790	0.757	0.302
C-12	-0.612	-0.619	-0.620	-0.591	-0.474	-0.197
O-13	-0.475	-0.475	-0.475	-0.469	-0.470	-0.303
H-14	0.149	0.149	0.150	0.142	0.111	0.072
H-15	0.141	0.143	0.143	0.138	0.110	0.062
H-16	0.148	0.152	0.152	0.145	0.112	0.082
N-17	-0.875	-0.875	-0.869	-0.844	-0.840	-0.420
H-18	0.364	0.363	0.361	0.348	0.345	0.217
H-19	0.307	0.299	0.296	0.293	0.306	0.184
HO-10	0.325	0.322	0.320	0.329	0.331	0.193
Dipole <sup>f</sup>	3.835	3.846	3.852	3.853	3.868	
Rms <sup>g</sup>	0.710	0.588	0.510	0.963	0.167	

<sup>a</sup> Geometry from reference 30.<sup>b</sup> Total number of points per molecule.<sup>c</sup> Number of points per atom located between van der Waal's [0, 3].<sup>d</sup> Number of points per atom located between van der Waal's [1, 3].<sup>e</sup> Values calculated using the program GAMESS.<sup>f</sup> SCF - dipole = 3.852647 D.<sup>g</sup> Root mean square deviations between quantum mechanically calculated and point charge derived electrostatic potentials in kcal/mol. See reference 24.

## References

1. R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833 (1955).
2. P. -O. Löwdin, *J. Chem. Phys.*, **18**, 365 (1950).
3. G. G. Hall, in *Advances in Atomic and Molecular Physics*, D. Bates and B. Bederson, Eds., Academic Press, New York, 1985, p. 41.
4. V. H. Smith, Jr., *Physica Scripta*, **15**, 147 (1977).
5. R. F. W. Bader, P. M. Beddall, and P. E. Cade, *J. Am. Chem. Soc.*, **93**, 3095 (1971).
6. P. Coppens, T. N. Guru Row, P. Leung, E. D. Stevens, P. J. Becker, and Y. W. Yang, *Acta Crystallogr. Sect. A*, **35**, 63 (1979).
7. D. A. Pearlman, and S. -H. Kim, *Biopolymers*, **24**, 327 (1985).
8. M. Gussoni, *J. Mol. Struct.*, **115**, 303 (1984).
9. S. Fliszár, *Charge Distributions and Chemical Effects*, Springer-Verlag, New York, 1983.
10. F. A. Momany, *J. Phys. Chem.*, **82**, 592 (1978).
11. S. R. Cox, and D. E. Williams, *J. Comp. Chem.*, **2**, 304 (1981).
12. U. C. Singh, and P. A. Kollman, *J. Comp. Chem.*, **5**, 129 (1984).
13. L. E. Chirlian, and M. M. Francl, *J. Comp. Chem.*, **8**, 894 (1987).
14. D. E. Williams, *J. Comp. Chem.*, **9**, 745 (1988).
15. D. E. Williams and J. -M. Yan, in *Advances in Atomic and Molecular Physics*, D. Bates and B. Bederson, Eds., Academic Press, New York, 1988, p. 87.
16. L. E. Chirlian, and M. M. Francl, *QCPE Bulletin*, **7**, 524 (1987).
17. For an example of the application of Lagrange multipliers see, R. Weinstock, *Calculus of Variations*, Dover, New York, 1974.
18. M. Dupuis, D. Spangler, and J. J. Wendoloski, National Resource for Computations in Chemistry Software Catalog, University of California, Berkeley, CA (1980), Program QG01. The present version of GAMESS is described in the Quantum Chemistry Program Exchange newsletter: M. W. Schmidt, J. A. Boatz, K. K. Baldridge, S. Koseki, M. S. Gordon, S. T. Elbert, and B. Lam, *QCPE Bulletin*, **7**, 115 (1987).
19. J. S. Binkley, M. Frisch, R. Krishnan, D. J. DeFrees, H. B. Schlegel, R. Whiteside, E. Fluder, R. Seeger, and J. A. Pople, *Gaussian 82*, Carnegie-Mellon University, 1982.
20. W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.*, **51**, 2657 (1969).
21. Since, for any given seed value, the same array of points will be generated, adding more points without changing the seed value will not result in a random distribution of the points and an artificial convergence of the charge values will occur.
22. PDQP makes use of a random number algorithm taken from: W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, New York, 1986, p. 196.

23. P. C. Hariharan and J. A. Pople, *Theoret. Chim. Acta*, **28**, 213 (1973).
24. The root mean square deviation between the quantum mechanically calculated and point charge derived electrostatic potentials is defined as:  $\text{RMS} = [N_p^{-1}Y]^{1/2}$ . It should be noted that the Fortran source code for CHELP (see reference 16) contains the following incorrect algorithm for the evaluation of the RMS deviation;  $\text{RMS} = N_p^{-1}[Y]^{1/2}$ . Hence, the values for the RMS reported in reference 13 are incorrect.
25. J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.*, **102**, 939 (1980).
26. T. Clark, J. Chandrasekhar, J. W. Spitznagel, and P. von R. Schleyer, *J. Comp. Chem.*, **4**, 294 (1983).
27. S. Takagi and G. A. Jeffrey, *Acta Crystallogr.*, **B33**, 3510 (1977).
28. R. J. Woods, K. A. Watson, K. Trouton, S. Fortier, and W. A. Szarek, *Carbohydr. Res.*, **193**, 1 (1989).
29. M. S. Lehmann, T. F. Koetzle, and W. C. Hamilton, *J. Am. Chem. Soc.*, **94**, 2657 (1972).
30. M. N. Frey, M. S. Lehmann, T. F. Koetzle, and W. C. Hamilton, *Acta Crystallogr.*, **B29**, 876 (1973).
31. R. J. Woods, V. H. Smith, Jr., W. A. Szarek, and A. Farazdel, *J. Chem. Soc., Chem. Commun.*, 937 (1987).