Potential Derived Charges Using a Geodesic Point Selection Scheme

MARK A. SPACKMAN

Department of Chemistry, University of New England, Armidale 2351, Australia

Received 22 November 1994; accepted 11 April 1995

ABSTRACT

Potential derived (PD) atomic charges, obtained by fitting to molecular electrostatic potentials, are widely used in molecular modeling and simulation calculations. These charges are known to depend on the sample of points chosen for the fit, on the particular point selection algorithm, on molecular translations and rotations in many instances, and even on molecular conformation. Following a critique of currently available methods, a novel point selection scheme is described which results in a highly isotropic array of points located on a series of fused-sphere van der Waals surfaces. The pattern of points is based on tesselations of the icosahedron, and these are discussed in some detail along with their connection with virus morphology, geodesic domes, and symmetric fullerene structures. Using methanol as a test case, it is shown that the new method leads to PD charges which are independent of translation and display minimal rotational dependence, and are hence far better suited to the determination of PD charges from electrostatic potentials obtained from both theory and experimental X-ray diffraction data. The conformation dependence of the newly derived PD charges for alanyl dipeptide is found to be substantially less than obtained earlier by Williams [*Biopolymers* **29**, 1367 (1990)].

Introduction

Molecular modeling and simulation techniques are increasingly effective and common in chemical research, especially in descriptions of intermolecular interactions and bulk phenomena. The accurate description of the molecular electrostatic potential is central to the success of these methods, and to date there have been two major approaches to estimating this important property: optimization and parameterization to reproduce experimental thermodynamic and structural data on fluids,[1,2] or using the results of *ab initio* calculations.[3,4] Typically, the interaction

CCC 0192-8651 / 96 / 010001-18

energy is expressed in the form

$$E = \sum_{\text{bonds}} k_r(r - r_e)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_e)^2 + \frac{1}{2}\sum_{\text{dihedrals}} V_n(1 + \cos(n\phi - \gamma)) + \sum_{\text{nonbounded}} \left(A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6} + q_i q_j/\varepsilon r_{ij}\right)$$

where the intermolecular electrostatic energy is derived from a point charge model of the molecular electrostatic potential. More sophisticated models, such as bond dipole or distributed multipole, are recognized as important improvements but are presently used infrequently; the point charge model is still widely recognized as a simple and fast approach to the computation of electrostatic interactions between molecules, and the efficiency of such computations remains a major factor in simulation studies.[5]

Kollman and co-workers have pioneered the use of point charges derived by fitting to the molecular electrostatic potential obtained from an *ab initio* calculation,[6] usually at the self-consistent field (SCF) 6-31G* level. These atomic charge models, in conjunction with other force field parameters, are capable of reproducing experimental differences in solvation energies to within 0.5 kcal mol^{-1}.[7] An important attribute of 6-31G* potential derived (PD) charges is that they yield sensible multipole moments and typically overestimate dipole moments by about 20%, and hence tend to compensate for the omission of molecular polarization in the foregoing energy expression.[3,4] A similar property can be ascribed to the OPLS charges of Jorgensen,[1,2] obtained as they are by fitting to experimental data on fluids.

Recently, the electrostatic potential derived from experimental charge density analysis of X-ray diffraction data has been used to yield PD charges for ammonium dimethylphosphate,[8] *N*-acetyl-α,β-dehydrophenylalanine methylamide,[9] and benzene and the COO^- groups in D,L-histidine and L-alanine.[10] At the very least, point charges derived in this manner offer the possibility of informing molecular simulation methods about realistic charges to use for molecules in the bulk and, perhaps in the future, may themselves be used directly in simulation studies. However, before these goals can be considered seriously, it is important to recognize that PD charges obtained from theory depend strongly on the way in which they were determined, as well as on molecular translations and rotations and even molecular conformation. Because of this, the experimental PD charges must be obtained in a manner which is essentially the same as that used to obtain theoretical charges, especially if direct comparisons are to be made between the two sets of charges.

In this article we present a novel point-selection scheme for use in deriving PD charges from fits to both experimental and theoretical electrostatic potentials. It is based on tesselations of the icosahedron, an approach which has been widely used in other areas of science, and one which results in a more isotropic set of points on a series of fused-sphere van der Waals surfaces around a molecule. In the next section we briefly discuss the least-squares determination of PD charges and then summarize other methods of selecting points with a special focus on that used by Kollman and co-workers. We then describe icosahedral tesselations in some detail and apply the new method to examine the rotational dependence of PD charges for methanol and the conformational dependence of PD charges for alanyl dipeptide, as well as more routine applications to several small molecules, making comparisons with previous work whenever possible.

Fitting Point Charges to the Electrostatic Potential

Regardless of the particular set of points selected for fitting, the least-squares determination of PD charges is common to all methods used and has been described in considerable detail elsewhere.[11-16] Here we summarize the concepts and terms employed in our later discussion. We seek to minimize the residual

$$\varepsilon = \sum_{\text{sites},\, i}^{N_p} \left[V_i - \sum_{\text{atoms},\, j}^{N_a} q_j/r_{ij} \right]^2$$

where V_i is the theoretical or experimental potential determined at N_p points, and q_j the charges located at the N_a nuclear sites, $N_p \gg N_a$. The set of PD charges which minimizes ε is unique for any particular set of N_p points and is obtained by linear least squares

$$\mathbf{Q} = \mathbf{A}^{-1}\mathbf{X}$$

where the vector/matrix elements are given by $Q_k = q_k$, $A_{kl} = \sum_{\text{sites},i}^{N_p} 1/r_{ik}r_{il}$, and $X_l = \sum_{\text{sites},i}^{N_p} V_i/r_{il}$. The goodness of the fit to the data is conventionally measured by two indicators, the root-mean-square (rms) fit, given by

$$\text{rms} = [\varepsilon/N_p]^{1/2}$$

and the relative rms derivation, usually expressed as a percentage, %rms, given by

$$\%\text{rms} = 100\left[\varepsilon \Big/ \sum_{\text{sites},i}^{N_p} (V_i)^2\right]^{1/2}$$

It should be clear from these expressions that the rms fit must be expressed in units, typically kcal mol^{-1}, while %rms is dimensionless. What is not so clear is that the rms fit is strongly dependent on the precise region around the molecule employed for the fit; points at greater distances have very small values of electric potential, and hence ε is generally very small, whereas regions close to the van der Waals surface commonly exhibit high potentials and consequently large rms fits. For this reason, erroneous conclusions may be drawn about the relative merits of fits to the potential using several different methods, if these conclusions rely solely on the rms fits. Unfortunately, although a relative figure of merit, the %rms fit also displays these characteristics because the long-range molecular electrostatic potential is easier to fit than that closer to the molecular surface (since at large distances the electrostatic nature of a molecule is readily described by a set of multipole moments). Perhaps a better indication of the quality of a set of PD charges is the set of estimated standard deviations (esd's) obtainable directly from the inverse least-squares matrix

$$\sigma(q_k) = \left[\varepsilon/(N_p - N_a - N_c)\right]^{1/2}(A^{-1})_{kk}$$

where N_c, the number of constraints imposed on the fit, is typically 1 (total charge constraint) but may be greater if the charges are constrained to fit the molecular dipole or higher moments. Williams often reports esd's in his work on PD charges,[17-19] but few other determinations of PD charges do so, although it is clearly a trivial additional effort. However, it is not at all clear what physical meaning may be directly ascribed to esd's determined in this manner, except that they are a measure of the curvature of the least-squares error surface at that particular coordinate. We provide evidence that under certain circumstances these esd's provide a good estimate of the variation of the PD charges with molecular (or coordinate system) rotation.

Summary of Point Selection Schemes

To date, many strategies have been used for the derivation of a set of points for determination of PD charges from *ab initio* and semiempirical electrostatic potentials. Two broad categories can be recognized—grid based and surface based—as well as methods using randomly generated points around the molecule.[14,16]

GRID-BASED METHODS

The original method of this kind is due to Momany,[20] and similar treatments have since been described by Cox and Williams,[21,22] Kroon-Batenburg and van Duijneveldt,[23] and Breneman and Wiberg.[24] All of these methods are based on the computation of the theoretical electrostatic potential on a Cartesian grid of points beyond the fused-sphere van der Waals surface of the molecule. They differ in specific atomic radii used as well as grid resolution (typically about 0.5 Å or less) and cutoff criteria at large distances. These methods are very easy to implement but inherently translation and rotation dependent, although this can be reduced by increasing the number and resolution of points.

SURFACE-BASED METHODS

Several surface-based methods have been described and utilized, and these overcome the transation dependence of the grid-based methods (unless the origin of the grid is fixed with respect to the molecular frame) but still display a rotational variation. The most commonly used methods are those due to Kollman and co-workers[6,13,25] (which are based on Connolly's molecular surface algorithm[26]) and Chirlian and Francl[11] (which uses a template of 14 points, in the directions of vertices and face centers of a cube, on the surface of a sphere). Other surface-based methods generate a set of random points on a spherical surface or in spherical shells,[14] utilize a spherical grid defined by equal increments in spherical polar angles,[27] or generate points on surfaces of constant electron density.[28]

Of these two broad categories, surface-based methods are the most attractive for our purposes, since we wish to use the same set of points to compute theoretical and experimental electrostatic potentials. This is not quite as trivial as it may seem, because the theoretical computation and charge density analysis of X-ray diffraction data invariably use different coordinate systems and orientations of the particular molecule, even if the molecular geometry is identical in the two studies. For this reason, we also wish to minimize the rotational dependence of the PD charges, and in this context it is worthwhile to discuss the drawbacks of the various surface-based point selection schemes (hence the mot ivation behind the present work). We can quickly dismiss the use of surfaces of constant electron density,[28] because this prevents the use of the same surface in both theoretical and experimental studies; the studies will always yield different electron density functions. Random point selection[14] is also unlikely to be of use for our purposes because it is fundamentally nonreproducible unless exceptionally large numbers of points are used. This appears to have been appreciated by Woods and co-workers,[14] who recommended the use of between 1500 and 2000 points *per atom* to ensure rotational invariance, an unworkably large number of points for large molecules. We demonstrate later that this is a considerable overestimate.

We are left with three possibilities: Connolly's algorithm,[26] the use of spherical templates with 14 points on the surface[11] (as used in CHELP), or a spherical polar grid.[27] The first is the most attractive, and we discuss it in considerable detail later. The CHELP choice of 14 points on a unit sphere has been widely discussed in the literature and is known to be inadequate, even when supplemented by additional points or multiple surfaces at different radii. This is a consequence of the sparse sampling and dependence of point location on the molecular coordinate system.[24,27] Nevertheless, the CHELP method is still widely used, even in discussions of conformational dependence of PD charges,[12] an application to which it would seem ill suited. Westbrook et al.[27] have used a spherical template based on lines of latitude and longitude separated by 15° in the respective polar angles, resulting in 266 points per unit sphere. This is clearly an improvement over the CHELP scheme, but a regular spherical polar grid will always result in an anisotropic distribution of points; the polar regions have a greater density of points than areas near the equator, and this must result in an inherent rotational dependence. We sought a point selection method which would yield as close to an isotropic point distribution as possible on the surface of a sphere. In the following section we discuss the Connolly algorithm, widely used by Kollman and co-workers, and then introduce a method based on tesselations of the icosahedron.

Connolly's Algorithm

As mentioned earlier, Kollman et al. have popularized the use of the surface generation algorithm of Connolly[26,29] and applied the resulting PD charges in numerous simulation studies.[3,4,6,13,25] The method is now readily accessible in the GAUSSIAN series of programs, in which the PD charges are referred to as Merz-Kollman/Singh or MKS charges[30]; it is also available in MOPAC93.[31] At first sight the method may appear to use Connolly's description of the smoothed van der Waals surface,[26] but it actually uses the dot surface numerical algorithm described briefly in Appendix II of that article; and the probe sphere, used to define a smoothed molecular surface by rolling over the van der Waals surface, in fact has a radius set to zero in the applications coded in GAUSSIAN and MOPAC93. This would also seem to be the algorithm used by Kollman et al. although, to our knowledge, it is never explicitly mentioned anywhere. Thus, the procedure used by Kollman and co-workers consists of using Connolly's dot surface algorithm, which produces approximately equally spaced points at a specified density on the surface of a sphere, with appropriate scaling to several multiples of van der Waals radii, translation to nuclear sites, and deletion of points inside neighboring spheres. The surfaces are the simple fused-sphere surfaces used by CHELP and many other methods.

Connolly's dot surface algorithm is generally claimed to produce points with a specified density (typically 1 $Å^{-2}$) on the scaled spherical surfaces, and thus the assumption is that any arbitrary point density can be used. Careful study of the algorithm shows this not to be the case; the algorithm possesses an inherent granularity, producing sets of points in a step-function manner, and will always yield a point density less than or equal to that specified. Typically, the output density of points is between 10 and 20% less than that requested, and the discrepancy is unpredictable and depends on how close to a step the requested

number of points lies. This granularity has actually been observed and commented on in other contexts. Meyer[32] has used Connolly's dot surface approach to compute molecular surface areas and observed that when the probe radius is gradually varied, the surface area does not change monotonically; rather it exhibits sharp peaks like a spectrum. The numerical peaks described by Meyer can all be rationalized with knowledge of the discrete nature of Connolly's spherical template algorithm.* A similar observation was made by Pascual-Ahuir and Silla[33] in comparing surface areas obtained with their GEPOL code with those from Connolly's algorithm.

In Figure 1 we demonstrate pictorially the granularity of Connolly's algorithm by plotting some of the resulting meshes of points (spherical templates) for $N \geq 20$.† Below each diagram we give the range of requested values of N which produce these templates (N is obtained from the requested surface density of points and the particular scaled van der Waals radius of the atom). It is readily apparent that the patterns of points arising from Connolly's algorithm display important anisotropies; the method utilizes spherical polar coordinates and hence the x and z directions are special and the xz and xy planes are mirror planes. Nevertheless, the density of points is nearly isotropic, especially for larger N. Typical values of N used in PD charge determinations for organic molecules, with a surface density of 1 $Å^{-2}$, are those less than 100.

* For probe radii of 0.14, 0.15, 0.16, and 0.17 (nm) and surface point density of 5 $Å^{-2}$, Meyer reports surface areas of 1.27, 1.34, 1.38, and 1.29 (nm^2) for norbornane. These values correlate strongly with the ratios of requested point density to actual point density for these four cases: 0.862, 0.915, 0.938, and 0.862.

† The actual granularity is somewhat finer than suggested by Figure 1. The allowed values of N up to 500 are 12, 14, 20, 23, 30, 34, 44, 49, 60, 64, 80, 85, 100, 106, 126, 129, 150, 156, 178, 187, 212, 218, 246, 255, 280, 290, 320, 331, 360, 370, 407, 417, 450, 460.

Geodesic Templates

An alternative to Connolly's method is to decorate a sphere with points based on tesselations of the faces of regular Platonic solids projected onto the spherical surface. In the following discussion we use the icosahedron because it has the highest symmetry, but its dual, the dodecahedron, would yield the same results. The method consists of tesselations of the triangular faces of the icosahedron with regular triangular or hexagonal grids. This simple technique has been described in varying amounts of detail elsewhere and used to discuss the morphology of viruses,[34-38] geodesic domes,[35,39,40] and, most recently, symmetric fullerenes.[39-48] Closely related methods have also been used in graphics rendering of molecular solids,[49] modeling lipid vesicles,[50] comparing surface structure,[51] estimating solvent exposure on protein surfaces,[52] and describing solvent accessible surfaces for self-consistent reaction field calculations.[33,53,54]

Triangular tesselations of the faces of the icosahedron were described in detail by Coxeter,[35] and subsequently by Fowler et al.[41,43] Hexagonal tesse-

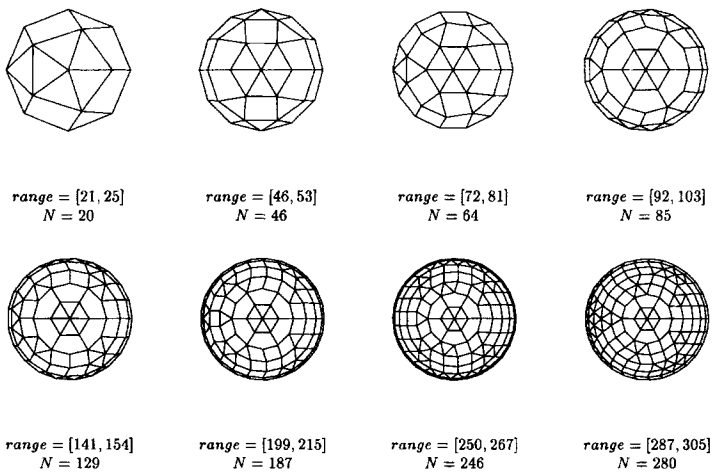

FIGURE 1. A selection of spherical templates obtained using Connolly's algorithm. The plane of projection, xy, is a mirror plane and the projection is down the z-axis. Not all possible templates are shown.

lations have also been treated by Fowler[43] and by Fujita et al.,[45] and similar ideas are evident in the work of Chiu et al.[48] Coxeter introduced the symbol $\{3,5+\}_{h,k}$ to denote the spherical tesselation described by the integers h and k on a triangular grid. An equilateral triangle is mapped commencing at one vertex of the grid and moving along h edges, and then changing direction by 60° and moving along k edges. Two further repetitions of this movement give a large triangle containing exactly $T = h^2 + hk + k^2$ smaller triangles. In Coxeter's symbol, the 3 refers to the number of sides per basic element of the grid and 5 + to the coordination of each vertex, at least five in the case of the triangular grid on the icosahedron. The duals of these spherical tesselations are given the symbol $[5+,3]_{h,k}$, which refers to objects with faces of five or more sides and three such faces meeting at each vertex, as first described by Goldberg.[55]

Most previous work has focused on the number of faces arising from constructions of this kind, but here we are interested primarily in the number of vertices because these form the template of points at which to compute the potential. We are also interested in maintaining maximum symmetry, point group I_h. As discussed by Fowler[41] for $h = k$ or $hk = 0$, this is the case, but in all other cases the rotational subgroup I results and there are two mirror images, $\{3,5+\}_{h,k}$ and $\{3,5+\}_{k,h}$. We have adapted the template construction method described by Lawrence et al.,[50] which determines a set of barycentric coordinates for the points on one triangular face and then applies this set to each of the faces in turn. The method is completely general and can be used to generate all $\{3,5+\}_{h,k}$ tesselations, but in the following discussion we use only those of I_h symmetry. For the hexagonal tesselations of this symmetry, we use the same templates and the identities $\{5+,3\}_{h,0} = \{3,5+\}_{h,h} - \{3,5+\}_{h,0}$ and $\{5+,3\}_{h,h} = \{3,5+\}_{3h,0} - \{3,5+\}_{h,h}$, which can be readily derived by inspection. Here, the subtraction indicates deletion of those points common to both templates. The resulting set of points is then projected onto the surface of a unit sphere and used as a template in the same manner as the Connolly templates.

Spherical templates for the first six members of each of the series $\{3,5+\}_{h,0}$, $\{3,5+\}_{h,h}$, $\{5+,3\}_{h,0}$, and $\{5+,3\}_{h,h}$ are illustrated in Figures 2 through 5, respectively. From these figures it is easy to see that the parent members of each series are the icosahedron, pentakisdodecahedron, dodecahedron, and truncated icosahedron, respectively, and the relationship of each template to a particular tesselation of the icosahedron is readily identifiable. These elegant and beautiful geodesic templates clearly represent highly isotropic arrays of points on the surface of a sphere, considerably more so than the arrangements resulting from Connolly's algorithm (Fig. 1). However, it is still

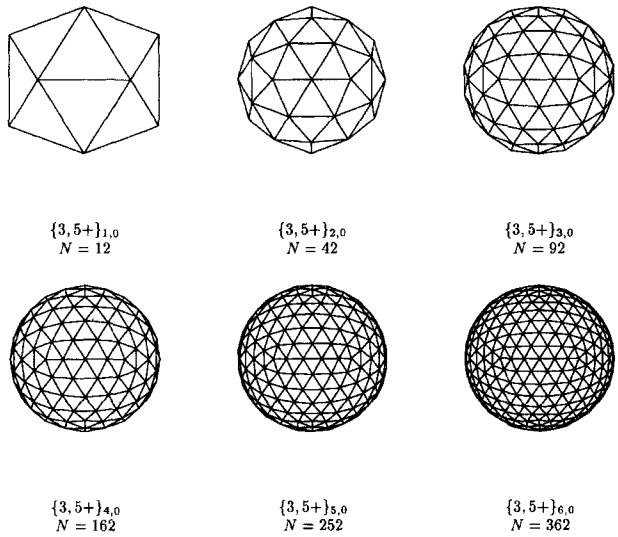

FIGURE 2. The first six geodesic templates of type $\{3,5+\}_{h,0}$ projected down a twofold axis. The plane of the page is a mirror plane.

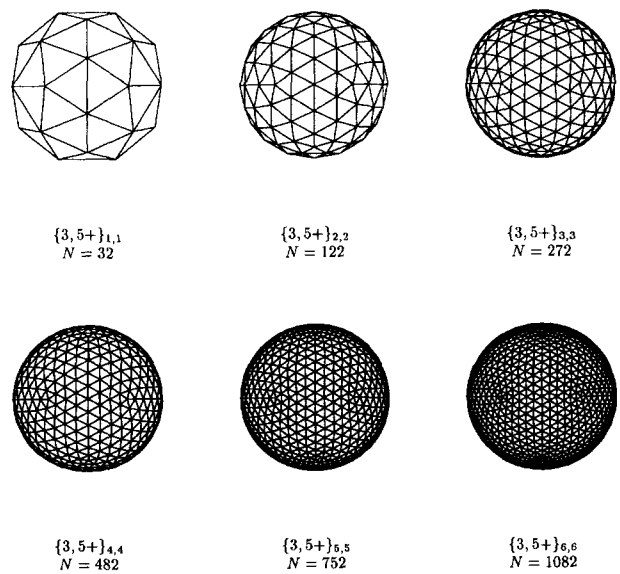

FIGURE 3. The first six geodesic templates of type $\{3,5+\}_{h,h}$ projected down a twofold axis. The plane of the page is a mirror plane.

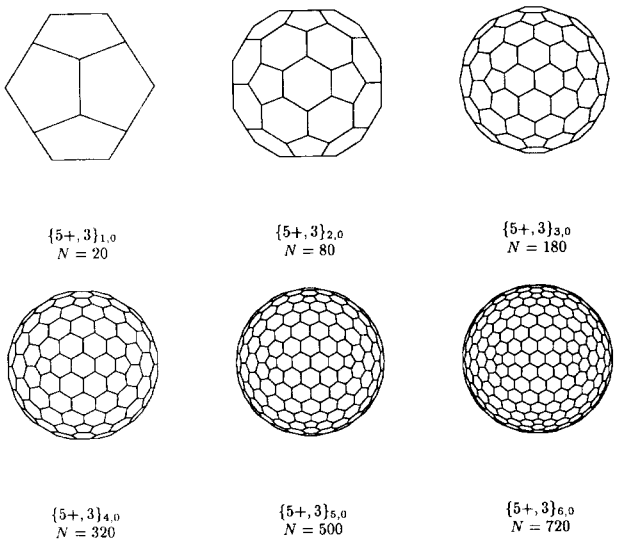

FIGURE 4. The first six geodesic templates of type $\{5+,3\}_{h,0}$ projected down a twofold axis. The plane of the page is a mirror plane.

evident that the point density is somewhat greater in regions near the vertices of the original icosahedron, a consequence of the gnomonic projection, which results in greater distortion near the center of the triangular faces. This is a subtle effect and unlikely to be important in the applications described later. Also given in these figures is N, the number of vertices for each template, and this information is summarized in Table I. All of these templates are possible candidates for use in the determination of PD charges, although, as will be shown in the next section, only those for which $N(h) < 400$ are likely to be of practical use in such work.

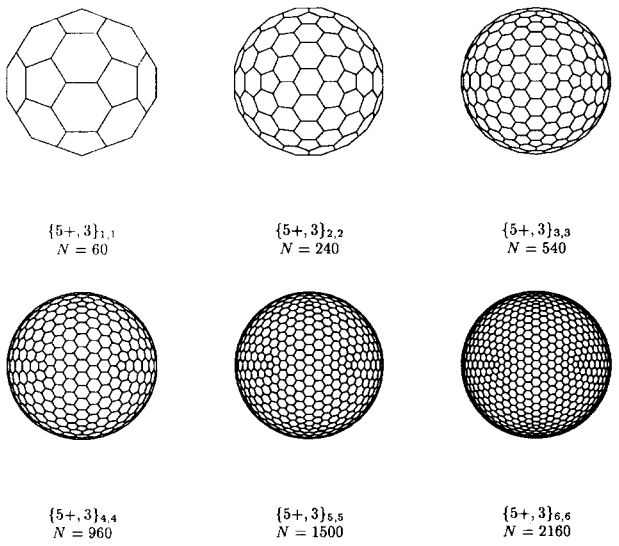

FIGURE 5. The fist six geodesic templates of type $\{5+,3\}_{h,h}$ projected down a twofold axis. The plane of the page is a mirror plane.

TABLE I.
Number of Points Generated by the Geodesic Templates in Figures 2 to 5.

Tesselation Type	$N(h)$	$h=1$	2	3	4	5	6
$\{3,5+\}_{h,0}$	$10h^2+2$	12	42	92	162	252	362
$\{3,5+\}_{h,h}$	$30h^2+2$	32	122	272	*482*	*752*	*1082*
$\{5+,3\}_{h,0}$	$20h^2$	20	80	180	320	*500*	*720*
$\{5+,3\}_{h,h}$	$60h^2$	60	240	*540*	*960*	*1500*	*2160*

Entries in italics are those which result in more than 3000 points for the methanol study.

Geodesic Versus Connolly Point-Selection Schemes

To assess the usefulness of the geodesic point-selection method in PD charge determination, in this section we provide a detailed comparison with the method used by Kollman et al. based on Connolly's surface algorithm. Three issues are the focus of this comparison:

1. Is there a difference in performance between the two methods, in terms of both the PD charges derived and their rotational and conformational dependence?
2. Is it possible to identify a threshold for either scheme (either points per geodesic template or point density on the surface) beyond which the results might reasonably be expected to be stable and well defined and little is gained by increasing the number of points?
3. What meaning may be attached to least-squares esd's for either scheme in this context?

To provide answers to these questions, we pursue three different avenues: (1) the rotational dependence of PD charges for methanol; (2) a direct comparison of PD charges obtained with both methods for a range of small molecules; and (3) conformational dependence of PD charges for alanyl dipeptide.

Rotational Dependence of PD Charges for Methanol

In this section we choose methanol as a single test case and examine the variation of PD charges from the two methods as a function of orientation of the molecule. The electrostatic potential is computed from an SCF/6-31G* wave function obtained at the SCF/6-31G* optimized geometry, and 10 random orientations of the molecule in space are chosen (actually 10 sets of three Euler angles chosen at random). To make the comparison as informative as possible, we compute the potential at points on the same four layers used by Kollman et al.—that is, at multiples of 1.4, 1.6, 1.8, and 2.0 times the van der Waals radii (1.20 Å for H, 1.50 Å for C, and 1.40 Å for O). We use the smallest 15 geodesic templates in Table I, spanning the range $95 < N_p < 2917$ for methanol. For the Connolly templates, we choose appropriate point densities which span the same range of total number of points; in the present case this corresponds to point densities up to approximately 6.7 $Å^{-2}$. For both methods we examine the mean charges and the range of charges for the 10 orientations as a function of N_p, especially in the context of the computed least-squares esd's. This detailed examination of rotational dependence has not previously been performed for the Connolly method, and we restrict this comparison to a single molecule, involving as it does calculations of the potential for 300 different sets of points; subsequent sections provide evidence that the conclusions are most likely general.

Figure 6 presents the range of charges obtained for each atom in methanol for both methods as well as the esd's in the PD charges; the esd is the same for both methods because it is strongly dependent on N_p when the selection of points is similar, as it is in the present case. Although the two H_s atoms (syn to the OH group) are symmetry equivalent, the results obtained in a numerical fit such as this will not in general be identical for both, and for this reason we present results for all six atoms in Figure 6. Note that in this figure the vertical axes are not all on the same scale, reflecting the different range of charges obtained for different atoms. For the hydroxyl H atom, the range is very small, and less than 0.015 e; for O the range is less than about 0.020 e; for the three methyl H atoms a slightly greater range is observed, but less than about 0.030 e; and for C the greatest range is observed, about 0.12 e. The significantly greater rotational variance of charges (and corresponding esd) for C in methanol has been observed by Cox and Williams[15,21] and by Woods et al.[14] and rationalized in terms of the location of C in the interior of the molecule, especially compared to O and H, and hence the inferior charge definition. This could be quantified by computing the rms distance of points used in the fit from each of the nuclear sites, but we do not pursue that possibility here.

Apart from different ranges in charges evident for different atoms in Figure 6, there are two other important observations to be made. Although the behavior for both methods is oscillatory (with decreasing amplitude as N_p increases), the range decreases in line with the esd, eventually becoming very close to the least-squares esd for both methods. The oscillations are important because they result from the granularity of the two point-selection algorithms; Connolly's method exhibits smaller oscillations for $N_p < 1000$, but for N_p greater than this the two methods show similar behavior. It would seem to be especially significant that for small numbers of points, less than about 1000 for methanol, large rotational variance can be expected with both methods; but as N_p increases beyond this, the esd is an excellent estimate of the rotational dependence of the charges for all atoms. Also worthy of note is that for methanol the esd initially decreases rapidly and then, beyond about 1500 points, decreases very slowly, in line with the $N_p^{-1/2}$ dependence in the defining equation.

Average PD charges for each of the six atoms are depicted in Figure 7 (i.e., the charge averaged over all 10 orientations of the molecule). Again, the vertical axes span different ranges (roughly the same as those in Fig. 6), but similar behavior is observed for all atoms. It is again easy to see the discrete nature of the two point-selection algorithms evident in the oscillatory behavior of the average charges as functions of N_p. In this instance, however, the geodesic method would appear superior to Connolly's method, because the oscillations are markedly less, and in fact the charges appear to converge rapidly in this case, especially for $N_p > 1000$. For the Connolly scheme, it is difficult to conclude that the mean charge will ever converge because significant oscillations are still noticeable for N_p near 2500.

To summarize the results on rotational dependence, we conclude that the geodesic and Connolly point selection schemes yield results which are almost, but not quite, indistinguishable: The

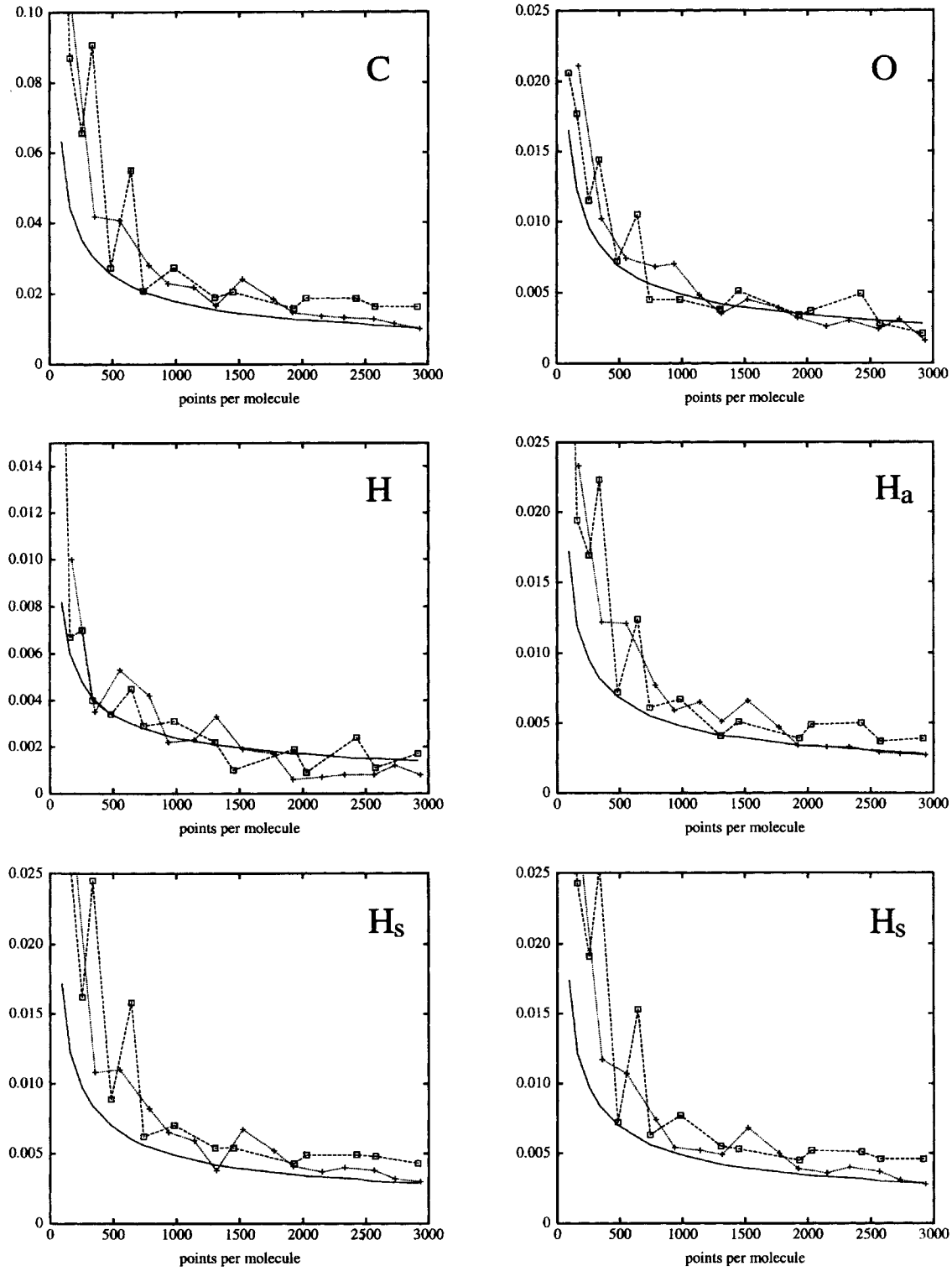

FIGURE 6. The range of PD charges (in electrons) for 10 random orientations of the methanol molecule. (---□---): geodesic templates; (··· + ···): Connolly's templates. The solid line is the least-squares esd, and is the same for both geodesic and Connolly templates.

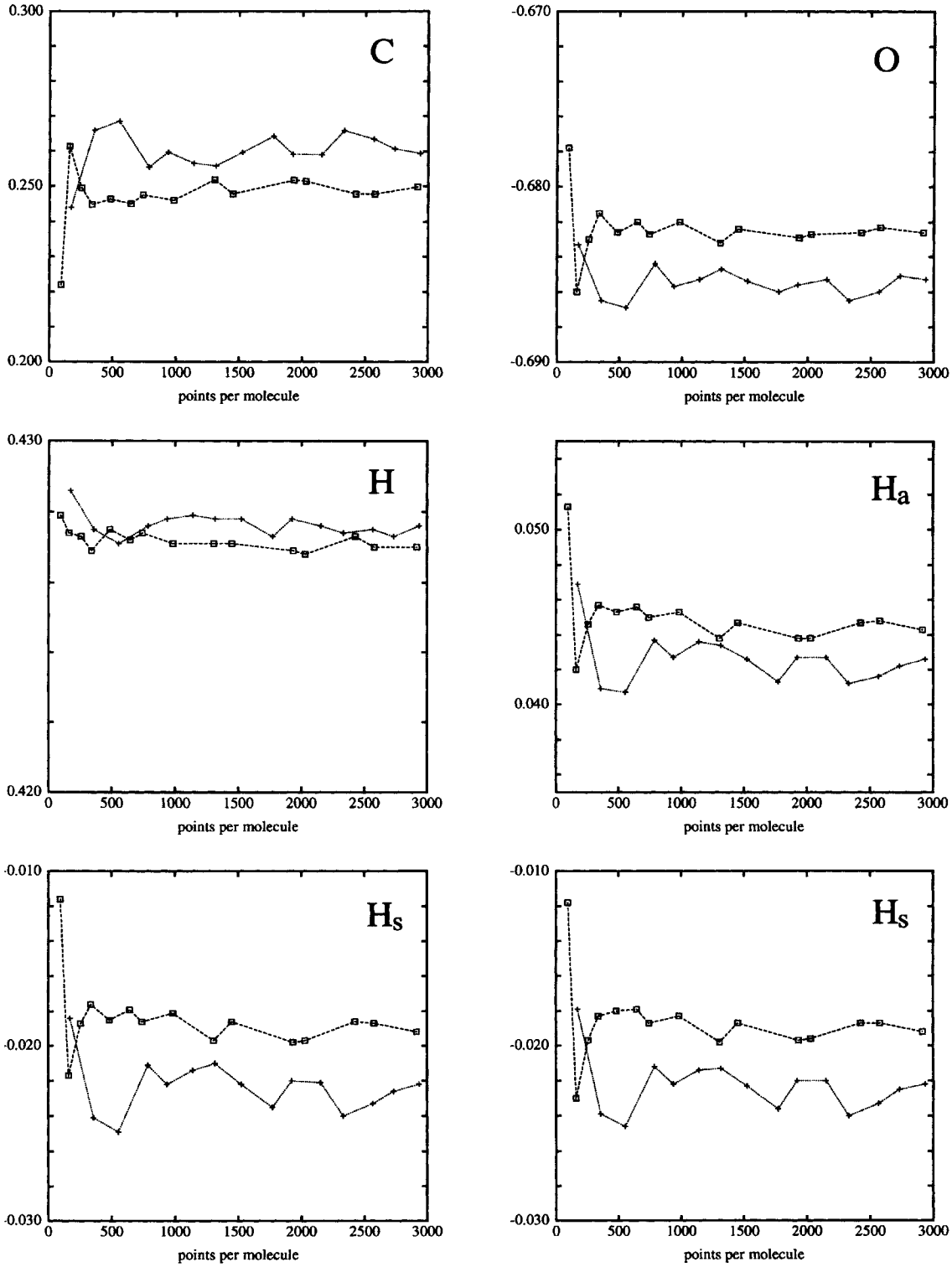

FIGURE 7. PD charges (in electrons) for methanol averaged over 10 random orientations of the molecule. (---□---): geodesic templates; (··· + ···): Connolly's templates.

least-squares esd's are the same for the same number of points and show the same behavior as N_p increases; both methods display oscillatory behavior in the range of charges, and this range decreases as N_p increases; the geodesic scheme displays significantly less variation in mean charges, while the Connolly scheme shows little sign of convergence in the mean charge even at 2500 points per molecule. For methanol, at approximately 1000 points per molecule, there would appear little to be gained (in terms of differences in charges or reduction of esds) by increasing N_p. This translates into the use of geodesic templates of the size of $\{3,5+\}_{3,0}$ or greater (see Table I), or point densities greater than 1.8 $Å^{-2}$ for Connolly's scheme. The latter point density is almost twice that routinely used with this point selection scheme,[13,25] although Cummins and Gready[56] recently reported using surface densities of at least 2 $Å^{-2}$ to obtain fitted dipole moments within 0.01 D of true values. Under the conditions described in this section, the least-squares esd's are best considered as a measure of the likely range of charge upon rotation of the molecule or coordinate frame; this rotational dependence can be reduced by increasing N_p, but it decays very slowly beyond a certain threshold.

One final comparison of the present results on rotational dependence is worth discussing. Breneman and Wiberg[24] proposed criteria for a grid-based method with dense sampling of points (CHELPG) which results in rotational variation of charges an order of magnitude less than that obtained with the CHELP method. They recommended a cubic grid of 0.3 Å spacing, and all points within 2.8 Å of a site are included in the fit. We have employed this recipe to determine PD charges for methanol, as well as ranges and esd's as a function of rotation in the manner described earlier. The dense sampling CHELPG method results in almost 5900 points for methanol, but even with this large value for N_p the variation of charges with orientation is generally greater than that observed in the present work for N_p near 3000. Specifically, the variations (in electrons) are 0.024 (C), 0.005 (O), 0.001 (H), 0.006 (H_a), and 0.007 (H_s). In contrast, the least-squares esd's are all correspondingly less than those shown on the far right of the graphs in Figure 6, essentially representing a continuation of the solid curves in those figures. As a consequence, the range of charges is typically three times the computed esd's for the CHELPG dense sampling scheme, reinforcing our conclusion that esd's can be regarded as estimates of rotational variance only under certain circumstances.

Although results for a single small molecule hardly constitute unassailable proof that one method is or is not generally superior to another, we believe that the results we have presented in this section demonstrate that the surface-based methods presented here are superior to grid-based methods and, moreover, that the geodesic point selection scheme represents an improvement over Connolly's method, while yielding charges which can be used in an identical manner to those reported by Kollman et al.[6,13,25] We recommend use of the geodesic point-selection scheme applied to four scaled van der Waals surfaces as described earlier, with two basic options: a $\{3,5+\}_{3,0}$ template or similar for routine work, and a $\{3,5+\}_{6,0}$ template or similar as an occasional check and for more demanding tasks.

PD Charges for a Range of Small Molecules

Table II summarizes results obtained with geodesic ($\{3,5+\}_{3,0}$ template) and Connolly (1.8 points $Å^{-2}$) point selection schemes for 14 small molecules at the SCF/6-31G** level. In almost all instances, the PD charges from the two schemes are identical within the respective esd's; exceptions are the methyl C atoms in CH_3CN and CH_3OH. These are typical "buried atoms" and, as seen in the previous section on rotational dependence, these atoms can be somewhat problematic—their charges are relatively poorly determined and susceptible to large changes by using only slightly different sets of points. This issue has been addressed in some detail by Stouch and Williams[57] and by Bayly et al.[58] In particular, the latter work nicely demonstrates that for methanol the least-squares error is relatively insensitive to the charge on the methyl carbon, but highly sensitive to the other charges [see Fig. 1 of that work; the curvature of the curves given in that figure is of course directly related to the esd's, $\sigma(q_k)$, reported in the present work]. Stouch and Williams suggested restraints on certain atomic charges to overcome these problems, and a procedure of that kind was implemented by Bayly and co-workers.[58] That aspect of PD charge determination is outside the scope of the present work, although we point out that the geodesic point-selection scheme recommended here would be a useful alternative to the Connolly scheme used by Bayly et al.[58]

Some further generalizations can be deduced from Table II: Both schemes faithfully reproduce

TABLE II.
PD Charges (in Millielectrons) for a Range of Small Molecules Obtained with Geodesic ($\{3,5+\}_{3,0}$ Template) and Connolly (1.8 points Å^{-2}) Point Selection Schemes.

Molecule		Geodesic	Connolly	Molecule		Geodesic	Connolly
HF	F	−458(2)	−457(2)	H_2O	O	−813(4)	−813(4)
	H	458(2)	457(2)		H	407(3)	407(3)
	N_p	444	486		N_p	528	556
	rms	1.5	1.4		rms	1.7	1.7
	%rms	10.0	9.9		%rms	11.4	11.1
	μ(1.97)	2.00	2.00		μ(2.20)	2.24	2.24
NH_3	N	−1101(7)	−1108(6)	CH_4	C	−498(6)	−489(5)
	H	367(3)	369(3)		H	124(2)	122(1)
	N_p	555	634		N_p	650	663
	rms	1.7	1.6		rms	0.4	0.3
	%rms	14.5	13.9		%rms	28.8	27.3
	μ(1.92)	1.96	1.97				
C_2H_2	C	−293(4)	−292(3)	C_2H_4	C	−330(9)	−338(9)
	H	293(1)	292(1)		H	165(3)	169(3)
	N_p	580	728		N_p	752	802
	rms	0.4	0.4		rms	0.9	0.8
	%rms	5.0	5.1		%rms	22.2	21.4
C_2H_6	C	−22(17)	−25(17)	CO_2	C	899(3)	897(3)
	H	7(4)	8(5)		O	−450(2)	−449(1)
	N_p	854	892		N_p	560	722
	rms	1.1	1.0		rms	0.7	0.6
	%rms	99.3	99.4		%rms	8.5	8.0
CH_2O	C	441(7)	449(6)	CH_3OH	C	266(20)	238(19)
	H	10(2)	7(2)		H	−24(5)	−17(5)
	O	−460(3)	−462(2)		H	42(5)	48(5)
	N_p	614	708		O	−688(6)	−680(6)
	rms	0.9	0.8		H	427(3)	428(3)
	%rms	6.6	6.2		N_p	758	789
	μ(2.67)	2.67	2.67		rms	1.4	1.3
					%rms	13.3	12.4
					μ(1.87)	1.86	1.85
$HCONH_2$	H	5(4)	0(4)	HCOOH	H	54(3)	48(3)
	C	717(13)	728(12)		C	696(9)	707(9)
	O	−593(4)	−595(4)		O	−575(3)	−577(3)
	N	−977(14)	−982(14)		O	−649(4)	−651(4)
	H	438(5)	440(4)		H	474(2)	474(2)
	H	409(5)	409(5)		N_p	655	789
	N_p	738	848		rms	1.1	1.0
	rms	1.2	1.2		%rms	8.7	9.0
	%rms	7.0	7.1		μ(1.60)	1.61	1.61
	μ(4.10)	4.11	4.10				
CH_3CN	C	−450(11)	−508(9)	CO_3^{2-}	C	811(6)	807(5)
	H	167(3)	182(2)		O	−270(2)	−269(2)
	C	460(4)	478(3)		N_p	640	815
	N	−519(2)	−517(1)		rms	0.9	0.8
	N_p	758	868		%rms	19.0	18.1
	rms	0.6	0.4				
	%rms	3.8	3.0				
	μ(4.04)	4.07	4.06				

Results refer to SCF / 6-31G** wave functions and optimized geometries. For dipolar molecules the *ab initio* dipole moment (in D) is given in parentheses.

the *ab initio* dipole moments; both schemes fail to provide a reasonable representation of the potential around hydrocarbons (in common with other methods, see Williams[18]); and the Connolly method usually, but not always, results in sightly lower rms and %rms fits to the potential, largely because it includes more points at greater distances (equal point densities on each scaled van der Waals surface), while the geodesic method results in relatively more points nearer the molecule (equal numbers of points on each scaled spherical template). Based on the results in Table II, it would be surprising if the two methods yielded significantly different results in simulation studies.

Conformational Dependence of PD Charges for Alanyl Dipeptide

There has been considerable attention given recently to the conformational dependence of PD charges.[12,17,19,24,57,59] Although that is not a concern of ours at present, we wish to demonstrate the performance of the geodesic point-selection scheme when applied to this subtle problem, having established the small rotational variance of the geodesic PD charges. We feel that a method which is not dependent on translations, and is minimally dependent on molecular rotations, is the only one capable of reaching meaningful conclusions about changes in PD charges with conformation. For this study we decided to duplicate the procedure followed by Williams in his study on alanyl dipeptide.[17] PD charges were computed using SCF/6-31G wave functions for the 12 different molecular conformers described by Williams; results were obtained with $\{5+,3\}_{1,1}$, $\{3,5+\}_{3,0}$ and $\{3,5+\}_{6,0}$ geodesic templates, and point densities of 1.0, 1.8, and 6.7 $Å^{-2}$ for the Connolly scheme, giving examples with roughly the same numbers of points for both schemes.

The atom numbering of the molecule (shown for conformer I) is given in Figure 8, and average charges and their range summarized in Tables III and IV. The PD charges using both surface-based point-selection schemes agree well with one another and are close to those reported by Williams, with the exception of the methyl group at C3A, for which a slightly greater charge separation is evident in Williams' results compared to the present ones (although the total charge on the methyl group differs little). Table IV reports the range of charges observed for the 12 different conformers with the three different methods. Here we can see a real difference between Williams' grid-based method and the two surface-based methods, as well as a subtle difference between the geodesic and Connolly schemes. The maximum range re-

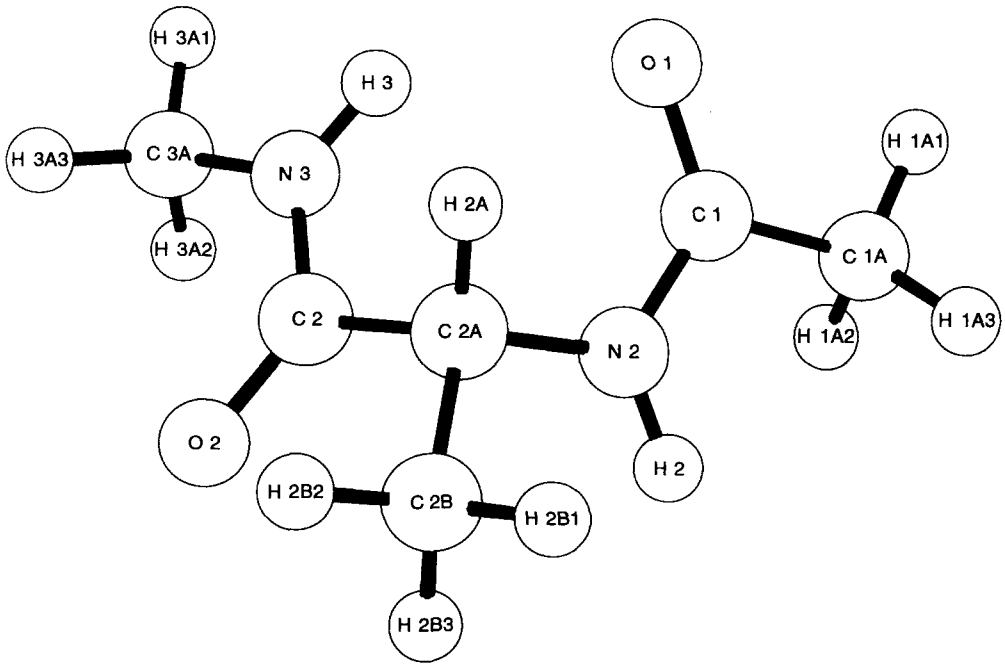

FIGURE 8. Atomic numbering for alanyl dipeptide (illustrated here for conformation 1).

TABLE III.
PD Charges (in Millielectrons) for Alanyl Dipeptide Averaged over 12 Conformations.

		Connolly Templates			Geodesic Templates		
		Points Å^{-2}					
	Williams	1.0	1.8	6.7	$\{5+,3\}_{1,1}$	$\{3,5+\}_{3,0}$	$\{3,5+\}_{6,0}$
Mean N_p	945	994	1889	7011	1159	1784	7019
C1A	−468	−510	−508	−503	−522	−519	−516
H1A1	149	156	157	155	160	160	160
H1A2	124	134	134	132	137	136	135
H1A3	126	133	132	130	135	134	133
C1	876	931	922	921	932	926	923
O1	−699	−709	−706	−707	−710	−709	−708
N2	−736	−781	−772	−772	−785	−777	−772
H2	348	368	367	367	370	368	368
C2A	271	198	186	184	230	217	202
H2A	51	76	80	79	71	73	77
C2B	−351	−313	−312	−307	−332	−324	−322
H2B1	87	76	77	75	80	79	79
H2B2	110	108	108	108	110	109	110
H2B3	105	102	101	100	106	104	103
C2	702	755	749	756	715	723	731
O2	−660	−671	−670	−671	−665	−667	−668
N3	−468	−547	−531	−541	−506	−508	−510
H3	344	342	337	339	331	331	331
C3A	−511	−237	−242	−227	−250	−249	−243
H3A1	191	132	132	129	134	133	132
H3A2	207	135	135	132	137	137	135
H3A3	204	124	124	121	124	123	122
rms / kcal mol^{-1}	1.1	1.0	1.0	1.0	1.1	1.1	1.1
%rms	3.0	6.7	6.7	6.8	7.1	7.1	7.2

Results are presented for the grid-based method of Williams,[17] three different surface point densities with Connolly templates, and three different geodesic templates corresponding to roughly the same total number of points as the three Connolly examples.

ported by Williams was 0.764 for C2A (the central sp^3 carbon atom), while the geodesic method results in a range of only 0.48 to 0.50 (depending on the template) for the same atom, and the Connolly scheme exhibits a range of 0.47 to 0.54 (depending on the point density). This difference between the methods is general with a few small exceptions, and on average the range exhibited by the geodesic or Connolly PD charges is almost 25% less than that reported by Williams. This is a considerable difference and one which surely has important implications for any attempt to derive meaningful conclusions about conformational dependence of PD charges: The conformation dependence can vary greatly between methods used to determine PD charges, and only a method with no translation dependence and minimal rotation dependence should be used in a study such as this.

This does not necessarily mean that the conformation dependence of PD charges reported in the literature is more imagined than real. In fact, the esd's given beside the ranges in Table IV show that the conformation dependences of PD charges for alanyl dipeptide appear to be highly significant and real for all atoms. This is especially clear from the 6.7 points Å^{-2} and $\{3,5+\}_{6,0}$ sets of charges in the table, where the esd's are very small and where for every atom the range far exceeds the least-squares esd. However, esd's can only indicate the significance of the differences between PD charges for various conformations given a specific method and point selection scheme. They cannot enlighten us about the more general question: Are such differences actually physically significant or just an artifact of the fitting procedure? Given that PD charges provide only a crude representation of

TABLE IV.
Range of charges (in Millielectrons) for 12 Conformations of Alanyl Dipeptide.

		Connolly Templates			Geodesic Templates		
		Points Å^{-2}					
	Williams	1.0	1.8	6.7	$\{5+,3\}_{1,1}$	$\{3,5+\}_{3,0}$	$\{3,5+\}_{6,0}$
Mean N_p	945	994	1889	7011	1159	1784	7019
C1A	148(27)	140(31)	66(22)	73(12)	123(29)	86(23)	59(12)
H1A1	38(7)	34(8)	20(6)	19(3)	27(7)	15(6)	18(3)
H1A2	59(7)	41(8)	29(6)	26(3)	35(7)	24(6)	23(3)
H1A3	50(7)	37(8)	32(6)	32(3)	38(7)	39(6)	32(3)
C1	197(18)	179(21)	195(15)	186(8)	199(20)	203(16)	198(8)
O1	78(5)	91(6)	92(4)	95(2)	94(6)	94(5)	(98(2)
N2	420(25)	314(26)	288(19)	280(10)	258(24)	268(20)	269(10)
H2	109(8)	89(8)	97(6)	97(3)	95(8)	99(6)	98(3)
C2A	764(33)	522(36)	474(25)	535(13)	499(32)	477(26)	494(13)
H2A	238(9)	177(10)	160(7)	168(4)	155(9)	159(7)	160(4)
C2B	295(29)	189(31)	200(22)	200(12)	216(28)	179(23)	195(11)
H2B1	102(8)	108(8)	104(6)	104(3)	97(7)	94(6)	100(3)
H2B2	116(8)	94(8)	95(6)	101(3)	95(7)	92(6)	96(3)
H2B3	140(8)	119(8)	122(6)	117(3)	121(8)	120(6)	121(3)
C2	254(21)	170(23)	229(16)	195(8)	220(22)	201(17)	203(9)
O2	80(5)	78(6)	96(4)	85(2)	85(6)	93(5)	89(2)
N3	176(20)	197(23)	192(17)	157(9)	182(22)	167(17)	162(9)
H3	90(7)	90(8)	88(6)	93(3)	99(7)	93(6)	94(3)
C3A	79(27)	165(32)	156(23)	143(12)	171(29)	128(23)	122(12)
H3A1	26(7)	66(8)	60(6)	61(3)	73(7)	60(6)	56(3)
H3A2	28(7)	48(8)	49(6)	44(3)	42(7)	36(6)	40(3)
H3A3	22(7)	21(8)	24(6)	21(3)	35(7)	25(6)	20(3)
Mean Range	160	135	130	129	135	125	125

Mean esd's are given in parentheses. Results are presented for the grid-based method of Williams,[17] three different surface point densities with Connolly templates, and three different geodesic templates corresponding to roughly the same total number of points as the three Connolly examples.

the molecular electrostatic potential, use of distributed multipole models in the future may make this problem irrelevant.

The difference between the geodesic and Connolly schemes in Tables III and IV is small but not insignificant. The mean range of charges (Table IV) decreases from 0.135 e to 0.129 e with increasing N_p with the Connolly method, and from 0.135 e to 0.125 e with increasing N_p for the geodesic scheme. Moreover, the geodesic scheme clearly converges more quickly and uniformly: The observed pattern of convergence of the average charges as a function of N_p (Table III) is monotonic for all atoms for the geodesic scheme, but often oscillates for the Connolly scheme (e.g., compare the two schemes for C2 and N3), and the mean range exhibited by the geodesic charges (0.125 e) is significantly lower than the Connolly result (0.129 e) for the largest N_p. The differences between the geodesic and Connolly schemes observed here are indeed small, but in complete agreement with the differences observed in the foregoing section on rotational dependence of PD charges for methanol. Together, they reinforce the conclusion that the geodesic point selection scheme offers an improvement over Connolly's scheme.

One further question remains: How important are these changes in PD charges with conformation? This is not the same as the question about their physical significance, but rather it is a query about the implications of conformational differences in PD charges on simulation studies. To attempt a partial answer to this, we examined how one set of mean charges ($\{3,5+\}_{3,0}$ geodesic results from Table III) reproduced the electrostatic potential for each conformer at the same set of points used for least-squares fitting in each case. Table V reports the results, in the form of a com-

TABLE V.
Comparison between Conformationally Averaged $\{3,5+\}_{3,0}$ Geodesic PD Charges and Those Obtained by Least-Squares Fitting for the 12 Conformers of Alanyl Dipeptide.

	rms / kcal mol^{-1}		%rms		μ / D	
Conformation	Best Fit	Average	Best Fit	Average	Best Fit	Average
1	1.14	1.54	8.6	11.7	3.21	3.03
2	1.08	1.64	8.1	12.4	2.97	3.14
3	1.18	2.16	8.3	15.2	4.53	3.93
4	1.11	1.57	7.0	10.0	5.34	5.71
5	0.98	1.90	5.5	10.7	7.09	7.55
6	1.16	3.07	6.0	16.0	7.72	8.72
7	0.98	2.11	5.8	12.4	6.94	7.48
8	1.10	1.61	7.6	11.1	3.94	4.13
9	1.04	2.07	5.3	10.5	8.12	8.76
10	1.15	3.39	5.9	17.5	7.84	9.03
11	1.13	1.41	8.8	11.1	0.34	0.48
12	1.08	1.48	8.2	11.3	1.76	1.89
Means:	1.09	2.00	7.1	12.5	4.98	5.32

Indices are given for the fit to the electrostatic potential (rms, %rms) as well as resultant dipole moments.

parison of rms and %rms deviations for the 12 conformers, as well as a comparison of dipole moments obtained from the fitted charges with those implied by the mean set of charges. The results clearly show that on average the rms deviation is increased from 1.1 kcal mol^{-1} to about 2.0 kcal mol^{-1} by using conformationally averaged charges; the %rms deviation also increases by a similar proportion; but the dipole moment is only slightly overestimated on average. This latter result is significant because it indicates that conformationally averaged charges are capable of reproducing molecular dipole moments for a variety of conformers of a floppy molecule, like alanyl dipeptide, to within about 0.35 Debye, despite the molecular dipole moments spanning a considerable range (the SCF dipole moments are within 0.05 D of those obtained from the PD charges obtained for each conformer, and they range from 0.32 D to 8.14 D). This seems a remarkable achievement and, although beyond the scope of this article, it would be extremely interesting to explore the use of such mean charges in simulation studies.

Conclusions

In the process of a careful consideration of the dual application of a PD charge determination procedure to both experimental and theoretical electrostatic potentials, we have arrived at an alternative point-selection method to those currently in use. We have characterized the geodesic method in some detail in this work, and we have recommended the use of geodesic templates of the size of $\{3,5+\}_{3,0}$ or larger for routine work. The geodesic templates clearly result in a more isotropic array of points on the surface of a sphere, perhaps the most isotropic array possible, apart from being elegant and beautiful in their own right. We have used them in the place of templates obtained from Connolly's dot surface algorithm and showed that the resulting charges are almost identical to those obtained by Kollman and co-workers in routine applications to a range of small molecules, yet there are small but significant advantages in the use of the geodesic method. In the process of testing our method, we have discovered that the least-squares esd's deriving from the geodesic and Connolly procedures can provide a good estimate of the rotational dependence of the PD charges, and this dependence can only be reduced slowly by increasing the total number of points beyond a certain threshold. In effect, we have attempted to quantify the rotational dependence of PD charges, and our conclusions suggest that recommendations based on some other methods, for use of as many as 3000 points per atom, are clearly unsupportable in general. Conformational dependence was also explored for alanyl dipeptide using the geodesic and Connolly methods and found to be significant (with respect to the esd's) for all atoms, although the magnitude was somewhat less than

reported previously by Williams. Again, the geodesic method was seen to represent a significant but small improvement over the Connolly point-selection scheme.

We have incorporated the geodesic point-selection method into the GAMESS[60] suite of programs, along with the CHELPG[24] and Kollman et al.[6, 13, 25] PD charge methods. In addition, the routines which generate the geodesic spherical templates are freely available from the author upon request.

Acknowledgments

The author is deeply appreciative of the opportunity to pursue the ideas reported in this work while on study leave in the Department of Chemistry at the University of Western Australia. This article is based on a talk given at Sagamore XI held in Brest, France in August 1994. This research has been supported by the Australian Research Council.

References

1. W. L. Jorgensen and J. Tiraldo-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).
2. W. L. Jorgensen, *Chemtracts Org. Chem.*, **4**, 91 (1991).
3. P. A. Kollman and K. M. Merz, *Acc. Chem. Res.*, **23**, 246 (1990).
4. P. Kollman, *Chem. Rev.*, **93**, 2395 (1993).
5. H. Sun, S. J. Mumby, J. R. Maple, and A. T. Hagler, *J. Am. Chem. Soc.*, **116**, 2978 (1994).
6. U. C. Singh and P. A. Kollman, *J. Comp. Chem.*, **5**, 129 (1984).
7. L. F. Kuyper, R. N. Hunter, D. Ashton, K. M. Merz, and P. A. Kollman, *J. Phys. Chem.*, **95**, 6661 (1991).
8. W. T. Klooster and B. M. Craven, *Biopolymers*, **32**, 1141 (1992).
9. N. Ghermani, N. Bouhmaida, and C. Lecomte, *Acta Cryst.*, **A49**, 781 (1993).
10. Z. Su and P. Coppens, *Z. Naturforsch.*, **48a**, 85 (1993).
11. L. E. Chirlian and M. M. Francl, *J. Comp. Chem.*, **8**, 894 (1987).
12. C. A. Reynolds, J. W. Essex, and W. G. Richards, *J. Am. Chem. Soc.*, **114**, 9075 (1992).
13. B. H. Besler, K. M. Merz, and P. A. Kollman, *J. Comp. Chem.*, **11**, 431 (1990).
14. R. J. Woods, M. Khalil, W. Pell, S. H. Moffat, and V. H. J. Smith, *J. Comp. Chem.*, **11**, 297 (1990).
15. D. E. Williams, In *Reviews in Computational Chemistry*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1991, Vol. 2, p. 219.
16. J. C. White and E. R. Davidson, *J. Mol. Struct.*, **282**, 19 (1993).
17. D. E. Williams, *Biopolymers*, **29**, 1367 (1990).
18. D. E. Williams, *J. Comp. Chem.*, **15**, 719 (1994).
19. T. R. Stouch and D. E. Williams, *J. Comp. Chem.*, **13**, 622 (1992).
20. F. A. Momany, *J. Phys. Chem.*, **82**, 592 (1978).
21. S. R. Cox and D. E. Williams, *J. Comp. Chem.*, **2**, 304 (1981).
22. D. E. Williams, *PDM93, Electric Potential-Derived Multipoles*, Department of Chemistry, University of Louisville, Louisville, KY, 1993.
23. L. M. J. Kroon-Batenburg and F. B. van Duijneveldt, *J. Phys. Chem.*, **90**, 5431 (1986).
24. C. M. Breneman and K. B. Wiberg, *J. Comp. Chem.*, **11**, 361 (1990).
25. K. M. Merz, *J. Comp. Chem.*, **13**, 749 (1992).
26. M. L. Connolly, *J. Appl. Cryst.*, **16**, 548 (1983).
27. J. D. Westbrook, R. M. Levy and K. Krogh-Jespersen, *J. Comp. Chem.*, **13**, 979 (1992).
28. J. Kendrick and M. Fox, *J. Mol. Graphics*, **9**, 182 (1991).
29. M. L. Connolly, *MS. Molecular Surface Program*, QCPE 429 (1981).
30. J. B. Foresman and A. Frisch, *Exploring Chemistry with Electronic Structure Methods. A Guide to Using Gaussian*, Gaussian Inc., Pittsburgh, 1993.
31. J. J. P. Stewart, *MOPAC 93.00 Manual*, Fujitsu Ltd., Tokyo, Japan, 1993.
32. A. Y. Meyer, *Chem. Soc. Rev.*, **15**, 449 (1986).
33. J. L. Pascual-Ahuir and E. Silla, *J. Comp. Chem.*, **11**, 1047 (1990).
34. D. L. D. Caspar and A. Klug, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1 (1962).
35. H. S. M. Coxeter, In *A Spectrum of Mathematics*, J. C. Butcher, Ed., Auckland University Press, Auckland, 1971, p. 98.
36. A. Klug, J. T. Finch, R. Leberman, and W. Longley, In *Principles of Biomolecular Organization*, G. E. W. Wolstenholme and M. O'Connor, Eds., J & A Churchill Ltd., London, 1966, p. 158.
37. R. W. Horne, *Virus Structure*, Academic Press, New York, 1974.
38. C. F. T. Mattern, In *The Molecular Biology of Animal Viruses*, D. P. Nayak, Ed., Marcel Dekker, New York, 1977, Vol. 1, p. 1.
39. R. Marks and R. B. Fuller, *The Dymaxion World of Buckminster Fuller*, Anchor Press/Doubleday, Garden City, NY, 1973.
40. T. Tarnai, *Philos. Trans. R. Soc. London, A*, **343**, 145 (1993).
41. P. W. Fowler, *Chem. Phys. Lett.*, **131**, 444 (1986).
42. D. J. Klein, W. A. Seitz, and T. G. Schmalz, *Nature*, **323**, 703 (1986).
43. P. W. Fowler, J. E. Cremona, and J. I. Steer, *Theor. Chim. Acta*, **73**, 1 (1988).
44. T. G. Schmalz, W. A. Seitz, D. J. Klein, and G. E. Hite, *J. Am. Chem. Soc.*, **110**, 1113 (1988).
45. M. Fujita, R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Phys. Rev. B*, **45**, 13834 (1992).
46. D. L. D. Caspar, *Philos. Trans. R. Soc. London, A*, **343**, 133 (1993).
47. J. S. Rutherford, *J. Math. Chem.*, **14**, 385 (1993).

48. Y. N. Chiu, P. Ganelin, X. Jiang, and B. C. Wang, *J. Mol. Struct.*, **312**, 215 (1994).
49. A. Dietrich and B. Maigret, *J. Mol. Graphics*, **9**, 85 (1991).
50. S. M. Lawrence, M. J. Lawrence, and D. J. Barlow, *J. Mol. Graphics*, **9**, 218 (1991).
51. P. L. Chau and P. M. Dean, *J. Mol. Graphics*, **5**, 97 (1987).
52. W. Kabsch and C. Sander, *Biopolymers*, **22**, 2577 (1983).
53. E. Silla, I. Tunon, and J. L. Pascual-Ahuir, *J. Comp. Chem.*, **12**, 1077 (1991).
54. A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 799 (1993).
55. M. Goldberg, *Tohoku Math. J.*, **43**, 104 (1937).
56. P. L. Cummins and J. E. Gready, *Chem. Phys. Lett.*, **225**, 11 (1994).
57. T. R. Stouch and D. E. Williams, *J. Comp. Chem.*, **14**, 858 (1993).
58. C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, *J. Phys. Chem.*, **97**, 10269 (1993).
59. J. J. Urban and G. R. Famini, *J. Comp. Chem.*, **14**, 353 (1993).
60. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, J. H. Jensen, S. Koseki, M. S. Gordon, K. A. Nguyen, T. L. Windus, and S. T. Elbert, GAMESS, QCPE Bulletin, **10**, 52 (1990).