J. Am. Chem. Soc. 1992, 114, 9075–9079

(0.43 mmol) was then added and the reaction allowed continue for 1.5 h. The reaction was then quenched with 5 mL of 0.1 M hydrochloric acid with noticeable evolution of gas, and the organic phase was extracted, dried over $MgSO_4$, and concentrated under vacuum. An NMR analysis of the solid organic residue revealed unreacted naphthalene and minor impurities of tin-coupled vinylic C=C*H*– resonances (<2%).

Acknowledgment. This research was generously supported by the National Science Foundation, Grant No. CHE 8304636. NMR equipment used in this research was financed in part with grants from the National Science Foundation to the Chemistry Department. Dr. Charles Cottrell, Central Campus Instrumentation Center, The Ohio State University, provided invaluable technical advice. We also acknowledge helpful discussions with Professor Theodore Cohen, University of Pittsburgh.

Atomic Charges for Variable Molecular Conformations

Christopher A. Reynolds,[†] Jonathan W. Essex,[‡] and W. Graham Richards*,[‡]

Contribution from the Department of Chemistry and Biological Chemistry, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, and Oxford Centre for Molecular Sciences and Physical Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QZ, United Kingdom. *Received August 28, 1991*

Abstract: The problem of generating high-quality atomic charges valid over a range of conformations has been addressed using two related methods which both employ a constrained minimization of the difference between the quantum mechanical and classical MEP (molecular electrostatic potential) with respect to the atomic charges. The first method involves determining the MEP and constraining the charges to reproduce the dipole at an alternative geometry. The second method involves determining the MEP for each conformation of interest and weighting the MEP for each conformation according to the appropriate Boltzmann factor. These methods offer considerable improvement over averaging the charges obtained at each conformation. The improvement in the performance of these multiple conformation MEP derived charges is illustrated by studying the variation of the classical dipole with conformation and comparing the results with those from ab initio calculations. It is proposed that the main use of these multiple conformation MEP derived charges and dipole constrained charges is likely to be in computer simulations where the ability to search conformational space is matched by the ability of the charges to yield the correct electrostatic properties at the conformations of interest. The errors arising from ignoring these effects have been assessed by evaluating the hydration free energy using a continuum method and are found to be significant. The extension of these methods to protein simulations is discussed.

Point charges on atoms are vital for many applications. A difficulty arises when the atomic point monopole charges change with molecular conformation. In particular, this is important in molecular dynamics. The ability to sample correctly alternative conformational regions is in some cases severely limited by the monopole model where the atomic charges are calculated at a single geometry but they are assumed to be valid for all geometries. Here we show that charges derived from the MEP (molecular electrostatic potential) determined for a range of conformations, with the MEP appropriately weighted according to the Boltzmann factor of the conformation, give a greatly improved description of the electrostatics of the molecule. These multiple conformation MEP derived charges yield superior free energy results when compared to standard single MEP charges in computer simulations of normal alcohols in water and carbon tetrachloride.[1,2]

The determination of reliable and theoretically justifiable atomic charges has long been a perplexing task for theoretical chemists since the description of atomic charges is shrouded in uncertainty, and yet at the same time the use of atomic charges in empirical potential energy functions provides an extremely powerful tool. Perhaps this usefulness is seen most strikingly in the recent application of free energy calculations to biological problems where calculated enzyme ligand binding energies in good agreement with experiment have been reported.[3] To date, many of these calculations have successfully used single MEP derived charges; but as techniques and computer power enable more extensive regions of conformational space to be sampled, the need will arise for charges valid over a wider range of conformations: a need which current methods cannot be guaranteed to meet.

Quantum mechanics offers the most general method of determining atomic charges, since other approaches such as fitting to experiment[4–6] are limited by the availability of experimental data. Moreover, different experimental properties may yield different values, particularly if the charges represent one part of a restricted force field. Thus if the force field does not treat polarization or if it treats repulsion inadequately, the charges are likely to be contaminated by polarization or repulsion effects. Moreover, if the experimental molecules have restricted mobility, as in a crystal, the resulting charges may only be valid for those restricted geometries.

Quantum mechanics does not provide a useful definition of an atom in a molecule. Bader partitioning[7] offers an elegant approach to the problem, but it is not a practical solution due to the computational expense. The Mulliken[8] and related population schemes are computationally inexpensive but do not yield reliable results. Extensions to the Mulliken scheme such as those developed by Huzinaga[9] and Stone[10] may offer more accurate results but are less useful in simulations because they introduce additional centers or higher order multipoles. Recently Hagler[11] has reported charges

---

[†] University of Essex.

[‡] Oxford University.

(1) Reynolds, C. A.; Essex, J. W.; Richards, W. G. *Chem. Phys. Lett.*, in press.

(2) Essex, J. W.; Reynolds, C. A.; Richards, W. G. *J. Am. Chem. Soc.* **1992**, *114*, 3634–3639.

(3) Reynolds, C. A.; King, P. M.; Richards, W. G. *Mol. Phys.* **1992**, *76*, 251–275.

(4) Hagler, A. T.; Huler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319–5327.

(5) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.

(6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(7) Bader, R. F. W. *Acc. Chem. Res.* **1985**, *18*, 9–15.

(8) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833–1840.

(9) Huzinaga, S.; Sakai, Y.; Miyoshi, E.; Narita, S. *J. Chem. Phys.* **1990**, *93*, 3319–3325.

(10) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233–239.

0002-7863/92/1514-9075$03.00/0

for planar systems based on the Cartesian derivatives of the molecular dipole. These charges compare favorably with MEP derived charges (see below), but their use in computer simulations is currently restricted to planar systems.[12] MEP derived charges[13-17] have proved popular since the MEP is rigorously defined;[18] the potential defines an interaction energy and so is relevant to molecular modeling, where interaction energies are calculated, and the potential has a classical analogue through which the atomic charges are defined by equating the classical and quantal MEP.

This work was developed as part of a program for calculating free energy differences in organic and aqueous solutions in order to calculate log *P*s "ab initio";[19] the calculation of log *P*s is discussed elsewhere.[2] This article describes two methods for determining atomic charges for small molecules. The main method uses multiple conformation MEPs; the other method uses a single MEP and dipole constraints to ensure that the dipole is correctly reproduced over several conformations. The article also discusses ways of applying these ideas to protein simulations.

Methods

MEP Derived Charges. The MEP upon which this work is based is defined as follows for both classical and quantum situations:

$$V(P)_{\text{quantal}} = \sum_{i=1}^{n} Z_i / r_{iP} - \sum_{p}\sum_{q} \langle \chi_p | 1/r_P | \chi_q \rangle \qquad (1)$$

$$V(P)_{\text{classical}} = \sum_{i=1}^{n} q_i / r_{iP} \qquad (2)$$

where Z_i is the nuclear charge, q_i is the atomic charge, r_{iP} is the distance from the nucleus to the point P where the potential is calculated, r_P is the distance from the nucleus to the element of charge density due to the atom-centered basis functions χ_p and χ_q, and n is the number of atoms. Since eq 2 is generally valid only beyond the van der Waals radius, $V(P)$ is compared at points on or beyond the surface of the molecule.

The best atomic charges can be defined as those which minimize the difference between $V(P)_{\text{classical}}$ and $V(P)_{\text{quantal}}$ subject to the constraint that the sum of the atomic charges, q_i, gives the molecular charge, q_{total}; this and other constraints are expressed using Lagrange multipliers, λ_l

$$\frac{\partial}{\partial q_i}(V(P)_{\text{quantal}} - V(P)_{\text{classical}})^2 + \sum_{l=1}^{m} \frac{\partial}{\partial q_i} \lambda_l \left(\sum_{i=1}^{n} q_i - q_{\text{total}}\right) = 0 \qquad (3)$$

$$\frac{\partial}{\partial \lambda} \lambda_l \left(\sum_{i=1}^{n} q_i - q_{\text{total}}\right) = 0 \qquad (4)$$

Equations 3 and 4 can be solved by the method of LU (lower–upper) decomposition and back substitution[21] using the following matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b} \qquad (5)$$

where **A** is a square matrix of dimension $n + m$ where n is the number of atoms and m is the number of constraints. (For molecule k, **A** is denoted $\mathbf{A}_k$). The potential related elements of **A**, **b**, and **x** are

$$A_k(i,j) = \sum_{P=1}^{N_P} 1/(r_{iP} r_{jP}) \qquad (6)$$

$$b_k(i) = \sum_{P=1}^{N_P} V_{k,P} / r_{iP} \qquad (7)$$

$$x_k(i) = q_i \qquad (8)$$

where N_P is the total number of MEP points, $V_{k,P}$ is the quantum mechanical electrostatic potential at point P for molecule k, and the additional elements are related to the constraints; the general form of the additional elements is given elsewhere.[17]

In the alcohol work, the ab initio quantum mechanical MEP and single MEP derived charges were calculated using CHELP[16,22] interfaced to GAUSSIAN 86.[23,24] The multiple conformation MEPs were also calculated on the CHELP surface[16] which consisted of about 500 points generated in concentric shells starting at 1.5 times the van der Waals radius.[20] The MEP, however, we calculated using GAUSSIAN 88,[25] and our own fitting program was used to derive the charges.[26] This program, *Rattler*, was also used in the calculation of the semiempirical multiple conformation MEP derived charges,[27] which were based on AM1[28] and MP3[29] wave functions calculated using MOPAC 5.0.[30]

In the threonine work, the 6-31G* calculations[31] used the direct SCF methods[32] implemented in GAUSSIAN 88, since threonine has 208 basis functions. The surface points were generated on four surfaces at 1.6, 1.8, 2.0, and 2.2 times the van der Waals radii.[20] An equal number of test points was used for each surface; this corresponded to a density of approximately 0.5 Å^{-2} on the 1.6 surface. Points falling within the appropriate surface of other atoms were discarded.

In this work we have chosen to use the 6-31G* basis set, because free energy calculations may be very sensitive to the basis set used to determine the charges,[33] with minimal basis sets giving very different answers from those from larger more flexible basis sets such as 6-31G*. (Many of the free energy perturbation calculations emanating from the Kollman group have used the 6-31G* basis set[34-39] even though the initial AMBER parameterization is based on the STO-3G basis set, since the 6-31G* MEP derived charges are considered to be more balanced vis-ã-vis the commonly used TIP3P water model.)[40] However, some of our calculations have also used the small split valence 3-21G basis set[41] as we have observed similar behavior to that of the 6-31G* basis set in terms of both free energies[33] and molecular similarities.[42]

Additional Constraints. It is possible to include additional constraints such as setting individual charges to a constant value (e.g., for the united-atom approach which sets the charges on apolar hydrogens to 0[5,43,44]), setting groups to have a constant charge (e.g., amino acid resi-

---

(11) Dinur, U.; Hagler, A. T. *J. Chem. Phys.* **1989**, *91*, 2949–2958.

(12) Dinur, U.; Hagler, A. T. *J. Chem. Phys.* **1989**, *91*, 2959–2970.

(13) Momany, F. A. *J. Phys. Chem.* **1978**, *82*, 592–601.

(14) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304–323.

(15) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.

(16) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894–905.

(17) Woods, R. J.; Khalil, M.; Pell, W.; Moffat, S. H.; Smith, V. H., Jr. *J. Comput. Chem.* **1990**, *11*, 297–310.

(18) Bonaccorsi, R.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1970**, *52*, 5270–5284.

(19) Essex, J. W.; Reynolds, C. A.; Richards, W. G. *J. Chem. Soc., Chem. Commun.* **1989**, 1152–1154.

(20) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

(21) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterline, W. T. *Numerical Recipes*; Cambridge University Press: Cambridge, 1989; pp 31–38.

(22) Chirlian, L. E.; Francl, M. M. *QCPE Bull.* **1987**, *7*, 39 QCPE 524; CHELP.

(23) Frisch, M. J.; Binkley, J. S.; Schlegel, H. B.; Ragavachari, K.; Melius, C. F.; Martin, R. L.; Stewart, J. J. P.; Bobrowicz, F. W.; Rohlfing, C. M.; Kahn, L. R.; Defrees, D. J.; Seeger, R.; Whiteside, R. A.; Fox, D. J.; Fleuder, E. M.; Pople, J. A. GAUSSIAN 86; Carnegie-Mellon Quantum Chemistry Publishing Unit: Pittsburgh, PA, 1986.

(24) Foley, C. K. *QCPE Bull.* **1988**, *8*, 67–69.

(25) Frisch, M. J.; Head-Gordon, M.; Schlegel, H. B.; Ragavachari, K.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. GAUSSIAN 88 Release C; Carnegie-Mellon Quantum Chemistry Publishing Unit: Pittsburgh, PA, 1988.

(26) Reynolds, C. A.; Ferenczy, G. G. *Rattler*; Oxford Molecular Limited: Magdalen Centre, Oxford Science Park, Stanford-on-Thames, Oxford OX4 4GA, United Kingdom.

(27) Ferenczy, G. G.; Reynolds, C. A.; Richards, W. G. *J. Comput. Chem.* **1990**, *11*, 159–169.

(28) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

(29) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221–264.

(30) Stewart, J. J. P. Mopac 5.0, Frank J. Seiler Research Laboratory, US Air Force Academy, Colorado Springs, Colorado 80840.

(31) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(32) Almloef, J.; Faegri, K., Jr.; Korsell, K. *J. Comput. Chem.* **1982**, *2*, 385–399.

(33) Lister, S. G.; Reynolds, C. A.; Richards, W. G. *Int. J. Quant. Chem.* **1992**, *41*, 293–310.

(34) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607–1614.

(35) Cieplak, P.; Bash, P.; Singh, U. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 6283–6289.

(36) Cieplak, P.; Singh, U. C.; Kollman, P. A. *Int. J. Quant. Chem.* **1987**, *QBS 14*, 65–74. Grootenhuis, P. D. J.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 4046–4051.

(37) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564–568.

(38) DeBolt, S. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 7515–7524.

(39) Grootenhuis, P. D. J.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 2152–2158.

(40) Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 5716–5720.

(41) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.

(42) Burt, C.; Reynolds, C. A.; Richards, W. G. To be submitted.

Table I. Comparison of the Propanol (6-31G*) Quantum Mechanical Dipole with the Dipole Calculated Classically Using 6-31G* MEP and Multiple MEP Charges for All Relevant Optimized Geometries[a]

	conformation about $O–C_\alpha$ and $C_\alpha–C_\beta$ bonds				
source of dipole	*aa*	*ag*	*ga*	*g+g–* (tight)	*g+g+* (loose)
μ_{quantal}	1.64	1.68	1.85	1.81	1.80
μ_q (multi MEP)	1.84	1.83	1.90	1.77	1.89
μ_q (*aa*)	**1.65**	2.23	2.48	2.33	2.90
μ_q (*ag*)	1.95	**1.64**	2.49	2.00	2.27
μ_q (*ga*)	2.12	2.47	**1.86**	1.95	2.23
μ_q (*g+g–*)	2.09	2.07	2.11	**1.82**	2.11
μ_q (*g+g+*)	2.40	2.22	1.94	1.79	**1.76**
	Population (293.15 K)				
	0.12	0.27	0.22	0.14	0.25

[a] Anti and gauche conformations are denoted *a* and *g*, respectively. Single MEP dipole moments evaluated at the geometry used to determine the MEP are shown in bold. The dipole moment is measured in debye.

dues within larger molecules may have an integral charge or carbonyl groups in amino acids may have equal and opposite charges as in ref 5), or setting the charges of similar atoms to be equal (e.g., the two N_η atoms in arginine). In this work, the threonine unit of the dipeptide was constrained to have a 0 net charge (see below). Generally, such constraints will yield reliable results, but if the constraints are too severe then the quality of the fitting may be poor, since the constraints are always reproduced exactly at the expense of the rms difference between $V(P)_{\text{quantal}}$ and $V(P)_{\text{classical}}$.

Dipole Constraint. The dipole moment may also be constrained to its experimental or quantum mechanical value by adding the following constraint

$$\sum_{i=1}^{n} c_j q_i - \mu_j = 0 \tag{9}$$

where c_j represents the *j*th component of the Cartesian coordinates and μ_j the *j*th component of the dipole moment, but since the MEP derived charges usually reproduce the quantum mechanical (and experimental) dipole well[15,16] this constraint is not usually necessary. Our studies have shown that while the MEP derived charges do indeed reproduce the quantum mechanical dipole at the geometry used in their determination (and in the determination of the $A(i,j)$ and $b(i)$ elements in eqs 6 and 7), denoted c_j, they may give severe errors in the dipole determined at some alternative geometry c_j'. Our first approach to this particular problem is to determine the MEP at the geometry denoted c_j and to include the dipole μ_j' for the alternative geometry c_j' as a constraint. This ensures that the dipole moment is reproduced closely for the first geometry, c_j, and exactly at the second geometry, c_j'.

Multiple Conformation MEP Derived Charges. The second approach to reproducing accurately the electrostatic properties over several conformations uses the fact that the form of the $A(i,j)$ and $b(i)$ terms in eqs 6 and 7 depends only on the number of atoms, not the number of points used to calculate $V(P)$. It is thus possible to weight these terms according to the Boltzmann populations, w_k, of each of the N_k conformations:

$$A(i,j) = \sum_{k=1}^{N_k} A_k(i,j) w_k \tag{10}$$

$$B(i) = \sum_{k=1}^{N_k} b_k(i) w_k \tag{11}$$

Solution of eq 5 thus gives charges valid for all conformations included in the sum. If some conformations are omitted, the charges may still be valid if either the missing conformations have a low weighting or other conformations contain the essential electrostatic features of the missing conformation. There are two approaches to determining the appropriate weighting for these multiple conformation MEP derived charges.

Table II. Comparison of the Propanol (3-21G) Quantum Mechanical Dipole with the Dipole Calculated Classically Using 3-21G MEP and Multiple MEP Charges, for All Relevant Optimized Geometries[a]

	geometry				
source of dipole	*aa*	*ag*	*ga*	*g+g–* (tight)	*g+g+* (loose)
μ_{quantal}	1.86	1.89	2.09	2.02	1.99
μ_q (multi MEP)	2.07	1.99	2.06	1.87	1.99
μ_q (*aa*)	**1.82**	2.37	2.56	2.39	2.99
μ_q (*ag*)	2.12	**1.85**	2.54	2.05	2.31
μ_q (*ga*)	2.21	2.55	**2.05**	2.17	2.41
μ_q (*g+g–*)	2.30	2.22	2.36	**1.98**	2.31
μ_q (*g+g–*)	2.52	2.35	2.13	1.95	**1.93**
	Population (293.15 K)				
	0.05	0.29	0.14	0.09	0.44

[a] Single MEP dipole moments evaluated at the geometry used to determine the MEP are shown in bold. The dipole moment is measured in debye.

Table III. Comparison of the Ethanol 3-21G and 6-31G* Quantum Mechanical Dipole with the Dipole Calculated Classically Using 6-31G* MEP and Multiple MEP Charges, for Both Optimized Geometries[a]

	3-21G basis set geometry		6-31G* basis set geometry	
source of dipole	*a*	*g*	*a*	*g*
μ_{quantal}	1.94	2.06	1.74	1.83
μ_q (multi MEP)	2.24	2.24	2.03	2.07
μ_q (*a*)	**1.93**	2.37	**1.73**	2.29
μ_q (*g*)	2.43	**2.35**	2.26	**1.83**
	Population (293.15 K)			
	0.28	0.72	0.38	0.62

[a] Single MEP dipole moments evaluated at the geometry used to determine the MEP are shown in bold. The dipole moment is measured in debye.

Table IV. Temperature Dependence of the Principal 6-31G* Multiple MEP Propanol Charges[a]

	10 K	293 K	1000 K
H	0.408	0.380	0.381
O	–0.644	–0.614	–0.616
C_α	0.189	0.203	0.203

[a] In units of the electronic charge.

Table V. Conformations Used To Determine Multiple MEP Charges for Threonine[a]

conformation	Φ	Ψ	Γ	weighting (unnormalized)
1	282.443	–30.877	62.192	0.178
2	282.443	–30.877	295.634	0.229
3	247.142	142.157	63.932	0.183
4	247.142	142.157	301.739	0.252
total				0.842

[a] Γ is defined by the dihedral angle $N–C_\alpha–C^\beta–O^\gamma$ and effectively defines the orientation of the C_γ atom; the angle Φ is approximately $\tau(C–N–C_\alpha–C_\beta) - 120°$.

1. Calculate the weights theoretically; for the alcohols these were obtained from self-consistent field calculations at the 3-21G or 6-31G* geometry. The weights were determined at 293.15 K, since the experimental log *P*s were obtained at this temperature. The weights are given by the standard Boltzmann formula using the energies of the optimized geometries. (A more rigorous treatment would have considered the shapes of these energy wells, but we do not believe this to be necessary at this stage.)

2. Use experimentally observed populations. This approach was used for the amino acid threonine. The populations of observed conformations were obtained by analysis of the entire Brookhaven Protein Databank,[45]

---

(43) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

(44) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.

(45) Brookhaven Databank, Brookhaven National Laboratory, Chemistry Department, Upton, New York 11973.

Table VI. (6-31G*) Quantum Mechanical Dipole of the Threonine Dipeptide and the Classical Dipole for the 6-31G* MEP and Multiple MEP Charges and Root Mean Square Fits between the Quantum Mechanical and Classical MEP, for Both the Single and Multiple Derived Charges[a]

	conformation											
	1a	1b	1c	2a	2b	2c	3a	3b	3c	4a	4b	4c
μ_{quantal}	7.487	5.773	7.283	5.735	7.752	7.824	2.529	3.547	1.542	2.659	3.753	1.626
μ_q (multi MEP)	7.305	5.752	7.574	5.721	8.092	8.258	2.731	3.766	1.508	2.858	3.913	1.723
μ (weighted charges)	7.739	5.922	8.064	6.023	*8.865*	*9.112*	3.312	4.383	1.757	3.193	4.292	1.664
μ_q (1a)	**7.477**	5.605	7.983	6.017	*8.940*	*9.385*	*3.685*	4.530	1.997	3.301	4.181	1.588
μ_q (1b)	7.528	**5.784**	7.828	6.005	8.555	8.729	3.477	*4.606*	2.355	3.606	4.513	2.023
μ_q (1c)	7.197	5.099	**7.266**	6.038	*9.031*	*9.223*	*4.195*	*5.292*	2.536	3.294	4.646	1.897
μ_q (2a)	7.355	5.270	7.559	**5.730**	*8.862*	*9.119*	*3.901*	*5.045*	2.012	3.272	4.699	1.846
μ_q (2b)	7.174	5.398	7.734	5.049	**7.754**	8.146	3.287	4.275	2.118	*3.663*	4.024	1.124
μ_q (2c)	7.471	5.712	7.892	5.038	7.664	**7.823**	3.043	4.340	2.292	*3.896*	4.391	1.627
μ_q (3a)	*8.651*	*6.933*	*9.041*	6.211	*9.008*	*9.135*	**2.542**	3.955	1.675	3.424	4.406	1.799
μ_q (3b)	8.227	6.607	*8.724*	6.251	*9.006*	*9.353*	2.555	**3.554**	1.134	2.699	3.889	1.708
μ_q (3c)	7.975	6.095	8.208	6.331	*9.449*	*9.664*	*3.573*	*4.641*	**1.547**	3.127	4.664	2.078
μ_q (4a)	8.305	6.394	*8.396*	*7.206*	*10.377*	*10.643*	*4.032*	*4.916*	1.484	**2.691**	4.726	2.431
μ_q (4b)	7.581	5.929	7.886	6.049	8.718	*8.995*	3.149	3.898	1.293	2.780	**3.792**	1.454
μ_q (4c)	7.999	6.285	*8.288*	6.185	*8.821*	*8.979*	2.956	4.154	1.823	3.184	4.189	**1.673**
rms[b]	0.445	0.511	0.472	0.548	0.564	0.504	0.567	0.515	0.524	0.500	0.626	0.575
rms[c]	1.138	1.069	1.491	1.193	1.241	1.194	1.109	1.219	1.053	1.377	1.004	0.942

[a] Single MEP dipole moments evaluated at the geometry used to determine the MEP are shown in bold. Classical dipoles differing from the quantum mechanical dipoles by more than 1.0 D are shown in italics. [b] Root mean square fit between quantum mechanical and classical MEP, for single MEP charges (in kcal mol^{-1}). [c] Root mean square fit between quantum mechanical and classical MEP, for multiple MEP charges (in kcal mol^{-1}).

in a manner similar to that of Ponder and Richards.[46]

Results

Tables I, II, and III show comparisons of the 6-31G* and 3-21G quantum mechanical dipole with the dipole calculated classically using the MEP and multiple conformation MEP charges and highlight the inadequacy of single MEP charges. Table IV illustrates the temperature dependence of the multiple conformation MEP charges for propanol.

The four main conformations representing 84% of all observed conformations of threonine are given in Table V, and these give rise to the 12 conformations used in Table VI when the orientation of the OH group is included (the H-atom is not observed in the crystal structure). Table VI shows a comparison of the quantum mechanical dipole with that determined classically in a similar manner to that given in Tables I–III.

Discussion

Dipole Analysis. An analysis of the conformational behavior of single and multiple conformation MEP derived charges can be found in Table I. This table shows the quantum mechanical dipole of propanol for all nine conformations resulting from rotations about the O–C_α and C_α–C_β bonds. The conformations are anti–anti, anti–gauche, gauche–anti, gauche–gauche(loose) (g_+g_+) and gauche–gauche(tight) (g_+g_-); four of these conformations are degenerate. The quantum mechanical dipole calculated at the optimized geometries varies little with conformation, from 1.64 to 1.85 D. The figures in bold show the classical dipole at the geometry used to determine the charges, indicating good agreement. However, the table also shows that the classical dipole varies from 1.79 to 2.90 D when the monopoles are used in an alternative geometry, indicating severe errors of over 1.0 D and 60%. The multiple conformation MEP monopoles, however, yield a classical dipole in good agreement with the quantum mechanical dipole at all geometries. The least satisfying agreement is for the anti–anti geometry where the difference is 0.2 D, but this conformation actually has a low Boltzmann weighting and, on average, the dipole is reproduced to within 0.1 D. A similar analysis applies to the 3-21G propanol dipoles, as shown in Table II; the corresponding effect for ethanol is not so marked as Table III shows. The weighting depends on the temperature, and so the charges are temperature dependent, as shown in Table IV. At high temperature there is little variation since all conformations have similar energies and hence the weights are given by their degeneracies.

Charges from Liquid Properties. The calculations reported in Tables I–III illustrate why the OPLS alcohol force field is reported to perform so well. Firstly, by careful design the dipole is conformationally independent since it has charges only on the first three atoms. Secondly, since the charges are obtained from liquid properties they automatically describe an average of the relevant conformations. Unfortunately, it is not possible to use liquid properties to obtain charges for all molecules of interest.

Protein Charges: Threonine Dipeptide. Threonine is possibly the nearest amino acid analogue of propanol and can adopt many conformations in a protein, having three back, bone torsional angles and two side-chain torsional angles. However, Ramachandran[47] style conformational probability maps show that the backbone conforms to certain well-characterized patterns, and similar patterns have also been observed for the side chains.[46] The analysis of side-chain orientations by Ponder and Richards considered only the internal residues of a restricted set of proteins. Since 1987, the Brookhaven databank has grown considerably and so we have reexamined the entire databank. The average conformations, corresponding to certain well-defined regions of conformational space, were determined and assigned weightings corresponding to their observed probabilities; the three unobserved orientations of the OH group were assigned equal weightings. The four most probable conformations shown in Table V, corresponding to 84% of all occurrences, thus yield 12 geometries which are similar to those in the rotamer library given in ref 46. Standard AMBER geometries were used to generate the coordinates, and the 6-31G* MEPs were generated from threonine containing $-COCH_3$ and $-NHCH_3$ at the N-terminus and C-terminus, respectively. Table VI shows the quantum mechanical dipoles, multiple conformation MEP derived classical dipoles, and the single MEP dipoles in a similar manner to Tables I, II, and III. The errors in the dipole exceed 2.8 D (65%). For these large molecules the error in the dipole is even greater than in propanol, and although the dipole is a rather crude indicator, it is clear that the use of multiple conformation MEP derived charges for flexible peptides will result in an improved performance over single MEP charges. Further evidence of this is given by the observation that peptide charges do indeed vary with conformation.[50] For proteins, where the conformation is known more precisely, restricted multiple con-

---

(46) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775–791.

(47) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283–437.

(48) Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304–5314.

(49) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 1712–1723.

(50) Williams, D. E. *Biopolymers* **1990**, *29*, 1367–1386.

formation MEP derived charges may be more appropriate.

The use of multiple conformation MEP derived charges obviously has a detrimental effect on the fitting process, but this is not very marked. The root mean square error between the classical and quantal MEPs is about 0.5 kcal mol^{-1}; this increases to about only 1.0 kcal mol^{-1} for the multiple conformation MEPs as shown in Table VI. The multiple conformation MEP charges are thus able to reproduce all the quantum mechanical MEPs remarkably well.

Large Systems. For large systems, ab initio calculation of all conformations may not always be feasible. In such circumstances it may be possible to use semiempirical methods to derive both the appropriate weightings and the charges, and to scale these charges so as to reproduce ab initio ones. We therefore determined semiempirical multiple conformation MEP derived charges as above, scaled according to the ratio of semiempirical to 6-31G* MEP charges at all geometries for ethanol and propanol respectively.[2] While the scaling approach[33] is common, its general use may require further investigation.[42]

Conformationally Averaged Charges. An alternative to weighting the MEPs is to weight the charges themselves. The dipole analysis for the weighted charges is shown in Table VI, row 3. The results are considerably worse than those obtained by weighting the MEPs. The maximum error in Table VI for the conformationally averaged charges (row 3) is 1.3 D whereas the maximum error for the multiple conformation MEP charges (row 2) is only 0.4 D. Moreover, since the MEP is rigorously defined and the charges are not, it is clearly preferable to weight the MEPs.

Assessing the Charges Using Free Energy Criteria. The analysis of the dipole variation shows that the single MEP charges are liable to severe errors but gives no indication of how these errors are expressed in terms of energy or free energy. We believe that such an assessment is important because we perceive that the multiple conformation charges may find their widest application in free energy difference calculations. For this purpose we have used the multiple conformation MEP derived charges as the standard against which the others can be assessed.

The errors in the free energy of hydration of propanol that would exist from the use of the single MEP charges in the wrong geometry, compared to those from the use of the multiple conformation charges, can be estimated using the Born equation for the free energy of solvation of a dipole centered in a sphere.[48,49] The error in the free energy of hydration of a sphere of dipole moment $\mu + \Delta\mu$, where μ is the true dipole, ϵ is the dielectric constant, and r is the radius is given by

$$\Delta G_{\mathrm{error}} = -\frac{2\mu\Delta\mu + \Delta\mu^2}{r^3}\frac{\epsilon - 1}{2\epsilon + 1} \qquad (12)$$

The radius may be estimated using

$$r = \pi/6\rho^{-1/3} \qquad (13)$$

where ρ is the number density for pure propanol. We assume that the propanol is entirely in the g_+g_+ conformation and that the error is due to the overestimation of the propanol dipole in this conformation. Using a density of 0.804 g cm^{-3}, a resultant radius of 4.94 bohr, $\epsilon = 80.0$, $\mu = 1.89$ D, and $\Delta\mu = 1.01$ D, $\Delta G_{\mathrm{error}}$ is -1.9 kcal mol^{-1}. Clearly this error is greater in water than in organic solvents. For carbon tetrachloride, $\epsilon = 2.228$, the corresponding error is -0.9 kcal mol^{-1}. These errors are significant and cannot be ignored in quantitative or semiquantitative work. Moreover, even though the Born equation is somewhat crude, the magnitude of the errors is comparable to that given by simulation methods.[1,2]

Conclusions

We have proposed two methods for generating atomic monopoles appropriate to multiple conformations. Both methods involve a constrained minimization of the difference between the quantum mechanical and classical MEP (molecular electrostatic potential) with respect to the atomic charges. The first method involves determining a single MEP and constraining the charges to reproduce the dipole at an alternative geometry. The second method involves determining the MEP at appropriate conformations and weighting the MEP for each conformation according to the appropriate Boltzmann factor. The improvement in the performance of these multiple conformation MEP derived charges is illustrated by comparing the classical dipole of the multiple conformation MEP and single MEP derived charges with the quantum mechanical dipole for all conformations.

The main use of these multiple conformation MEP derived charges is likely to be in Monte Carlo or molecular dynamics simulations where the ability of these methods to search conformational space is matched by the ability of the multiple conformation MEP derived charges to yield the correct electrostatic properties in these conformations. Since free energy appears to be sensitive to electrostatics, multiple conformation MEP derived charges are likely to yield superior results in free energy perturbation calculations; this has indeed been observed for the mutation of ethanol to propanol.[2] Multiple conformation MEP derived monopoles are also likely to be useful in protein simulations as the multiple conformation MEPs can easily be constructed according to the probabilities observed for naturally folded proteins.

Acknowledgment. We wish to thank G. G. Ferenczy for useful discussions; J.W.E. wishes to acknowledge the financial support of Glaxo Group Research through a Glaxo prize studentship.

Supplementary Material Available: Tables containing the AM1, PM3, 3-21G, and 6-31G* charges for all conformations of methanol, ethanol, and propanol and the AM1, PM3, and 3-21G Mulliken dipole analysis for all conformations of ethanol and propanol (5 pages). Ordering information is given on any current masthead page.