ELSEVIER

Chemical Physics 246 (1999) 37–47

Chemical Physics

www.elsevier.nl/locate/chemphys

Unusual distributed charge models of water's electric potential

Seiichiro Tanizaki [a], Janez Mavri [b], Harry Partridge [c], Peter C. Jordan [a,*]

[a] *Department of Chemistry, MS-015, Brandeis University, PO Box 9110, Waltham, MA 02454-9110, USA*
[b] *National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia*
[c] *NASA Ames Research Center, Moffett Field, CA 94035-1000, USA*

Received 28 October 1998

Abstract

Based on wave functions determined from high level quantum mechanical theory, we critically evaluate some distributed charge models for the electrical potential of a water molecule. In three cases we find that for optimal parameterization several charges coalesce to form a multipole, reducing to four site descriptions, each of three point charges and a multipolar site. In two cases, the multipole is located at the distributed charge; in the third it is sited at the oxygen atom. All three limiting models are considerably superior to simple four site point charge approximations. The description of water's electrical potential can be dramatically improved, even close to the H and O atoms, if distributed charges are described as diffuse spherical charge densities. Some models constructed in this way can admit of intuitive 'chemical' interpretations.

Keywords: Multipolar site; Wave functions; Oxygen atom

1. Introduction

The development of accurate, reliable force fields is a recurrent theme in computational chemistry. All efforts start by treating the molecular charge distribution as a limited set of point charges or multipoles, the particular choice governed by how the force field is to be used. In standard approaches to modeling liquid water [1–3], the charges (and their location) do not describe the isolated water molecule; they are adjusted in fitting behavior of the liquid. Such an approach relies, to some extent, on error compensation. As long as the mean atomic environment in a simulation differs sufficiently little from that used in establishing the force field parameters, standard potentials [2,4,5] have been used with great success. However, in many applications, especially those involving ionic and polar species in non-traditional environments (simulation of hydration at hydrophobic interfaces or of the properties of ion selective channels are premier examples), water's mean molecular charge distribution is very different from that in bulk water [6,7]. To model such behavior also requires accounting for polarization which affects both dynamics and thermodynamic averages [6–11]. In addition, quantum dynamical simulations of proton transfer processes require the use of polarizable solvent models. Electronic polarization can follow proton motion while orientational polarization cannot [12,13].

* Corresponding author

0301-0104/99/$ - see front matter
PII: S0301-0104(99)00158-5

Practical engineering of water force fields requires incorporating as much of the underlying chemistry in as realistic a way as possible. The first step in this process is to accurately model the molecule's equilibrium charge distribution [14,15]. To do this, one seeks an artfully chosen, but not necessarily unambiguous [16,17] set of point charges and/or multipoles, one that reproduces its electrostatic potential or its electric field nearly exactly. This problem has a long history; only a few relatively recent representative water papers are cited [18–22]. It should be noted that simply increasing the number of parameters is ultimately counterproductive in development of an effective representation. Recent work illustrates the nature of the engineering difficulties [23] and the importance of chemical insight [24]. We reopen this issue and use the results of recent highly accurate quantum calculations [25–27] to assess a number of such models. Previous investigations have used similar data to assess the reliability of well known computational potentials and to construct accurate, but computationally unwieldy, representations of the electrostatic potential; some representative examples include [19–21]. We seek distributions useful for a shell model analysis [28], an approach recently successfully applied to molecular nitrogen [29] and to water interaction with a MgO surface [30].

The shell model requires negative charges representing electrons in addition to the nuclei. A particular advantage to shell models is that, in keeping with reality, they lend themselves naturally to treating the 'electrons' as diffuse charge distributions rather than simply point charges. Their mean location is unspecified; analysis begins by considering the underlying distributed point charge models (DPCM). In water, the negative charges would represent bonding and lone pair electrons. Determining optimal sites for distributed charges is a non-linear least squares problem, considerably more complex than constructing an atomic charge model (ACM) description of the electric potential [31]. It is not always clear how locations of the distributed sites have been chosen [19,21]. Further, models which optimally fit the electrostatic potential are often inconsistent with chemical intuition. For example, in a well known DPCM for water the best picture had inverted lone pairs [32].

Section 2 describes the methods for generating and fitting the electrical potential and the model charge distributions tested. Section 3 presents the results of the various approximate electrical descriptions of water. We first reproduce earlier results: certain DPCM descriptions, while adequately portraying water's electrical properties, yield totally counterintuitive parameterizations [32]. We find that some optimal DPCMs limit naturally to distributed multipoles and determine three parameterizations of this type; each is well suited for development of a point polarizability model for water. They essentially meld the charge site picture of TIP4P [3] and the Barnes polarizable water [33] pictures yielding models with three atomic sites and one distributed site. In two the distributed site is multipolar, while in the third it is the oxygen site. We then show that by describing the shell charges as diffuse greatly improves some models' ability to describe water's electrical potential and that counterintuitive DPCM parameterizations can be the consequences of ignoring this feature. Section 4 summarizes our analysis.

2. Methods and models

The electrical potential is fit to the DPCMs subject to two sets of constraints to ensure that the potential limits properly far from the water molecule. In Case A, the only limitation is electroneutrality. In Case B, neutrality, dipole moment and quadrupole moment constraints are used. The parameters are determined by minimizing the target fitting function Ξ, where

$$\Xi = \sum_{j=1}^{N} w_j \left(V_j^0 - V_j^{\mathrm{cal}} \right)^2 . \qquad (1)$$

V_j^0 is water's electrical potential, determined from high accuracy quantum calculations and V_j^{cal} is the electrical potential of the model DPCM; w_j is a weighting factor, generally chosen to be unity, and potential differences are sampled at N points.

After explicitly incorporating constraints, Eq. (1) is minimized by the Levenberg–Marquardt method [34,35] with $w_j \equiv 1$. Goodness of fit is determined from two quantities related to Eq. (1), the rms devia-

tion of a test charge e_o interacting with the potential σ, and the relative rms deviation s:

$$\sigma = e_o[\,\Xi/N\,]^{1/2}; \tag{2}$$

$$s = \left[\,\Xi/\sum_j^N \left(V_j^0\right)^2\right]^{1/2}. \tag{3}$$

We examine seven increasingly complex models, denoted P3, P4, T4, P5, T5, P6 and T6, involving three to six charge sites; P-models are P(lanar) and T-models are T(hree dimensional). Except for T4, three charges are always located at the nuclei. The remaining charges are distributed to satisfy specific shell model scenarios. The charges and the locations of distributed charges are parameters; model characteristics are summarized below. Planar and three dimensional models are illustrated separately in Fig. 1 which also defines the coordinate system.

P3: a two parameter ACM, similar to SPC [2] or TIP3P [3]. P4: a four parameter DPCM, an extension of TIP4P [3] with a charged O-site. Here the fourth charge represents a mean bonding charge distribution, located on the molecule's symmetry axis between the O and the Hs. T4: a four parameter DPCM, related to ST2 [1]. Two charges are located at the Hs and two, expected to be on the opposite side of the O, represent individual lone pairs; the O site is uncharged. P5: a six parameter DPCM. The fourth and fifth charges represent the bonding charge distribution and the mean lone pair charge distribution, respectively. Both are located on the molecule's symmetry axis; they are expected to lie on opposite sides of the O. T5: a five parameter DPCM. This is the extension of T4 in which the O site is charged (or equivalently of P3 with explicit lone pairs). P6: a seven parameter DPCM. The fourth and fifth charges represent bonding charge distributions for the individual OH bonds; these are expected to lie (roughly) along the OH bonds. The sixth charge models the mean lone pair distribution, located as in P5. T6: a seven parameter DPCM. The fourth charge represents the bonding charge distribution, located as in P4. The fifth and sixth charges represent the individual lone pairs, located as in T4 and T5.

Results from two high level quantum calculations based on the experimental molecular geometry (OH distance 0.9572 Å, HOH angle 104.52°) were used to compute water's electrical potential, electric field

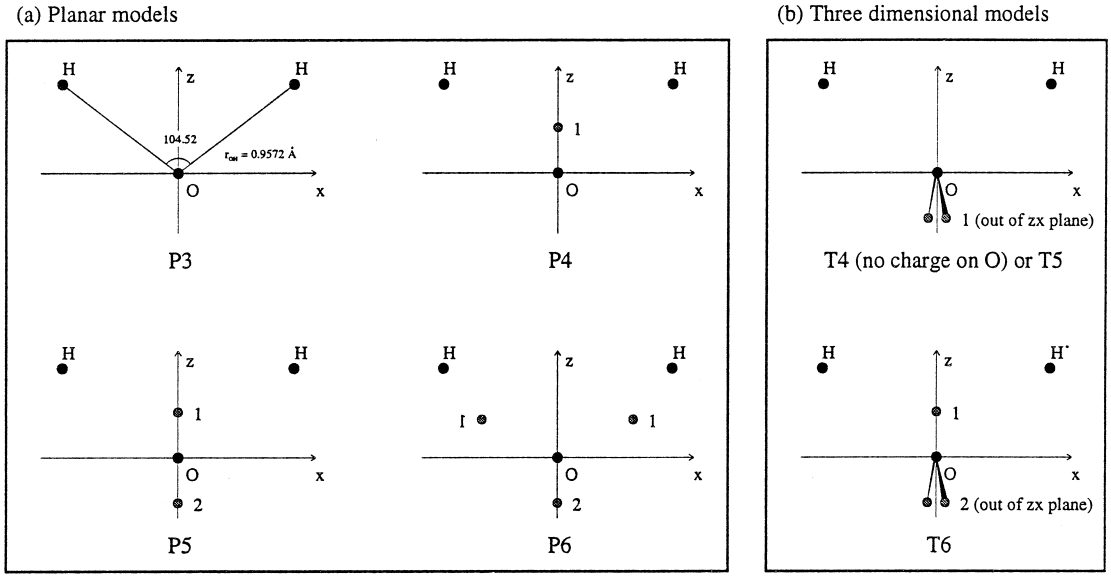

Fig. 1. Schematic diagrams of the various models studied, grouped into (a) planar and (b) three dimensional sets (see text). The *y*-axis is perpendicular to the molecular plane.

Table 1
Multipole moments, relative to center-of-mass, determined using K93t[a] and PS[b] basis sets (dipole moments in Debye, quadrupole moments in Buckinghams, octupole moments, Ω, in esu Å^3 and hexadecapole moments, Φ, in esu Å^4

Moment	Experimental[c]	K93	Error	PS	Error
μ_z	1.8546	1.8924	+2.0	1.8383	−0.9
Θ_{xx}	2.63	2.5555	−2.8	2.5696	−2.3
Θ_{yy}	−2.50	−2.4097	−3.6	−2.4368	−2.5
Ω_{zzz}		−1.3140			
Ω_{xxz}		2.1932			
Φ_{zzzz}		−1.3393			
Φ_{xxzz}		1.6215			
Φ_{xxxx}		−0.3652			

[a] Refs. [25,26].
[b] Ref. [27].
[c] Refs. [39,40].

and the various multipole moments. The first (K93) employed both extended and polarized basis sets, (12s8p3d/6s5p7d) for oxygen and (6s2p1d/4s2p1d) for hydrogen, with calculations performed at the fourth order Møller–Plesset level with single, double and quadruple excitations (MP4SDQ) [25,26], giving rise to 66 basis function for the water molecule; calculations of the wave functions and derived electrical properties were carried out on the Hewlett-Packard C180 workstation at the National Institute of Chemistry, Ljubljana. [1] The second (PS) used an aug cc-pV5Z basis set at the complete active space self-consistent field/multireference configuration interaction level [27] for determination of wave functions and electrical properties. Multipole moments (center-of-mass based), which are very sensitive to the level of theory, are summarized and compared with experimental data in Table 1. The two-wave functions are very similar. For model parameterization and to compare wave functions, electric potentials and fields were computed on nine evenly spaced, concentric, oxygen centered spheres with radii from 2.0 to 4.0 Å. Small radii were used since, ultimately, we plan to use the water model to describe hydrogen bonding and to treat aqueous interaction with small ions. Radii larger than 4 Å exceed the Lennard-Jones radius of the O; sampling at such distances provides no information beyond that already incorporated by imposing multipole moment constraints. Rms differences between the potential and field strength computed with the PS and K93 wave functions were ~2–3% on each sphere; the mean field orientation angle differences were ~0.7° on each sphere. Because higher multipole data could be easily extracted from the K93 wave function it was used in most parameterizations even though, based on differences between theory and experiment for the lower multipole moments, the PS data might have been slightly preferable. On individual spheres the potential was sampled by varying θ from 0° to 180° and ϕ from 0° to 90° all in increments of 5° for a total of 6003 sampling points, 667 on each sphere. Increasing the sampling intervals to 10° had no significant effect on the parameterizations, a well known result from ACM analyses [36,31]. Using different weighting factors w_j for points at different distances from the O also had little influence on the model parameters. All calculations were based on double precision Fortran programs.

3. Results

We now contrast results for the seven models under neutrality (A) and multipole moment (B) constraints. Tables 2 and 3 present the results. The quantities $\sigma_{1.6}$ are rms errors excluding points highly unlikely to be sampled in binary interaction between two waters or between an ion and water; the cut-off chosen is 1.6 Å from either H-atom, configurations unimportant for hydrogen bonding [37].

3.1. Distributed point charges

The more complex models monotonically reduce the errors. The most substantial change is between P3 and P4 where rms errors (σ or s) decrease threefold. Furthermore, Case B models (Table 3) describe the electrical potential slightly less well than Case A models (Table 2). The reverse is true for

[1] We thank Kersti Hermansson for providing us her basis set.

Table 2
DPCM parameters for Case A with fitting function Eq. (1) minimized subject to neutrality constraint only; σ is the rms deviation of the potential (in kJ mol^{-1}), s is the percent error of fit, Eqs. (2) and (3), $\sigma_{1.6}$ is the rms deviation (points closer than 1.6 Å to an H-atom excluded), and n_{param} is the number of model parameters. The multipolar entries are percent deviations of the moments' absolute values from the quantum calculation. Charges are in units of electronic charge (e_o) and distances in Å

n_{param}	P3 2	P4 4	T4 4	P5 ≡ DP4 6, 5[a]	T5 5	P6 7	T6 ≡ QP4 7, 6[b]	T6′ ≡ QP4$_O$ 7, 6[b]
q_O	−0.6980	1.7911	0.0	−0.2641	−0.2882	−4.9479	−2.8394	23450.7393
q_H	0.3490	0.5621	0.5251	0.5701	0.5129	0.5219	0.5214	0.4863
q_1		−2.9153		2161.4638		0.0719	336.1154	−8.2010
z_1		0.0882		0.5547714		−0.3315	−0.3417325	−0.1013
x_1						0.7849		
q_2			−0.5251	−2162.3399	−0.3688	3.7609	−167.1594	−11721.7555
z_2			0.2060	0.5547076	0.2748	−0.0443	−0.3429023	0.00004277
y_2			0.2469		0.3275		0.0161506	0.002351
σ	9.31	3.03	2.64	2.84	2.56	2.11	2.13	2.19
s	16.34	5.33	4.57	4.99	4.50	3.70	3.74	3.83
$\sigma_{1.6}$	7.27	2.74	2.37	2.60	2.32	1.67	1.68	1.78
μ	3.8	1.9	1.3	1.1	1.1	0.9	1.1	1.0
Θ_{xx}	−44.7	−8.7	−6.1	−6.7	−5.4	−2.2	−2.4	−1.8
Θ_{yy}	39.4	4.5	0.9	3.8	1.2	−1.2	−0.9	−1.1
Ω_{zzz}	23.1	−24.6	−11.8	−7.4	−6.7	−1.0	−5.9	−4.1
Ω_{xxz}	−22.4	25.3	19.0	22.0	17.7	5.2	6.3	7.4
Φ_{zzzz}	30.5	−12.1	−3.3	4.9	0.8	−6.0	−6.5	2.9
Φ_{xxzz}	−30.3	12.3	5.2	6.5	3.2	5.3	9.5	−2.8
Φ_{xxxx}	11.1	−43.2	−36.2	−31.0	−36.1	−17.8	−46.0	−24.1

[a] DP4 model.
[b] QP4 and QP4$_O$ models.

the higher multipole moments; in general, the more constrained models fit the higher multipoles better.

Except for P3 and P4, the optimal parameterizations have non-intuitive features, for T4 and T5 identical to those found previously (inverted 'lone pairs') [32]. None is suitable for shell model application since q_O is uniformly negative. The six parameter P5 is less satisfactory than the simpler T4 and T5 pictures; the planar description of water's charge distribution, even with more parameters, is inferior to either of the three dimensional ones. The number of free parameters, to be determined from water's electric potential, is less than n_{param}. In general for Case A, it is $n_{\text{param}} - 1$ and for Case B, $n_{\text{param}} - 4$.

In the optimal P5 the distributed charges limit counterinuitively, to a charge-dipole ($q_2 \sim -q_1 \gg e_o$ and $z_2 \sim z_1$); both are in the 'bonding electron' region and there is no indication of a 'lone pair' and no shell model interpretation. The limiting distribution is illustrated in Fig. 2a. The limiting parameters can be determined by a direct fit to a four site, five parameter model, denoted DP4 with a charge, $q = q_1 + q_2$, and a dipole located at z: [2]

Case A: $z = 0.5548$ Å, $q = -0.8760\ e_o$, $\mu = 1.0384$ D;

Case B: $z = 0.6611$ Å, $q = -0.8318\ e_o$, $\mu = 1.1485$ D.

The optimal P6 description is totally non-intuitive. All distributed charges are positive; the only negative charge is at the O site. The distributed charges are in 'non-physical' locations. The 'bonding' charges are not found in the bonding region $z_1 > 0$; in Case B the 'average lone pair' is found in a totally non-physical region $z_2 \gg z_H$. Unlike all

[2] The notation refers to a four site model, multipolar (up to Dipole) at one site. QPn models, discussed below, refer to n-site models, multipolar (up to Quadrupole) at one site.

Table 3
DPCM parameters for Case B with fitting function Eq. (1) minimized subject to both neutrality and multipole moment constraints

n_{param}	P3[a] 2	P4 4	T4 4	P5 ≡ DP4 6, 5[b]	T5 5	P6 7	T6 ≡ QP4 7, 6[c]	T6′ ≡ QP4$_O$ 7, 6[c]
q_O	−0.6724	0.8361	0.0	−0.3709	−0.4952	−2.9238	−3.4334	1037.8258
q_H	0.3362	0.6015	0.5655	0.6015	0.5197	0.4967	0.5125	0.4798
q_1		−2.0391		797.1371		0.9643	11748.5428	−4.5220
z_1		0.1524		0.6611881		−0.099126	−0.2954770	−0.1506
x_1						0.24951		
q_2			−0.5655	−797.9692	−0.2721	0.0018	−5873.0672	−517.1317
z_2			0.1906	0.6609435	0.3950	1.778405	−0.2955200	0.0008211
y_2			0.2376		0.4149		0.0029460	0.01161
σ	9.54	3.83	3.31	3.19	2.95	2.30	2.34	2.35
s	16.73	6.72	5.81	5.60	5.18	4.02	4.10	4.13
$\sigma_{1.6}$	7.00	3.17	2.79	2.91	2.65	1.66	1.64	1.70
Θ_{xx}	−46.7							
Θ_{yy}	41.6							
Ω_{zzz}	25.9	−35.2	−24.8	−5.8	−6.5	0.1	−3.8	0.0
Ω_{xxz}	−25.2	34.5	28.2	26.7	22.5	5.1	4.2	5.2
Φ_{zzzz}	33.0	−20.2	−11.9	13.7	3.5	6.0	−4.1	3.7
Φ_{xxzz}	−32.9	20.3	13.5	6.8	6.1	−2.8	6.9	−3.9
Φ_{xxxx}	14.4	−53.5	−46.2	−27.4	−45.2	−20.2	−41.9	−22.7

[a] Only the dipole moment can be constrained.
[b] DP4 model.
[c] QP4 and QP4$_O$ models.
Conventions of Table 2 apply.

other models, Case A and Case B parameterizations are qualitatively different.

Analysis of T6 yields two solutions, of roughly equal σ. In one the distributed charges are in close proximity, q_1 and q_2 of opposite sign and the O charge negative. This T6 representation is equivalent to a distributed multipolar site at $z \sim (z_1 \sim z_2) < 0$ where $q = q_1 + 2q_2$. In the alternate solution there is a shell charge at $z_1 < 0$ while the q_2 are close to O, forming a multipole. For the K93 wave function this is a local σ minimum; for the PS wave function it is the global minimum. The two parameterizations are equally attractive. Limiting parameters are found by direct fits to the four site, six parameter models, QP4 and QP$_O$, where, in both cases, $\Theta_{xx} = \Theta_{zz} = -\Theta_{yy}/2$. For QP4 the multipole is sited at z and while for QP4$_O$ it is at the O. Both are illustrated in Fig. 2b. For QP4 we find:

Case A: $z = -0.3420$ Å, $q = 1.7935\, e_o$,
$\mu = 1.9359$ D, $\Theta_{yy} = -0.4308$ B;

Case B: $z = -0.2952$ Å, $q = 2.4108\, e_o$,
$\mu = 2.4288$ D, $\Theta_{yy} = -0.4926$ B;

and for QP4$_O$ we find:

Case A: $q_o = 6.7309\, e_o$, $\mu_o = -4.6933$ D,
$\Theta_{o,yy} = -0.6214$ B,
$z_1 = -0.1045$ Å, $q_1 = -7.7039\, e_o$,
$q_H = 0.4865\, e_o$;

Case B: $q_o = 3.4935\, e_o$, $\mu_o = -4.0648$ D,
$\Theta_{o,yy} = -0.6684$ B,
$z_1 = -0.1522$ Å, $q_1 = -4.4536\, e_o$, $q_H = 0.4800\, e_o$.

For both models, the distributed source is in the 'lone pair' region. Both parameterizations are improvements over the simpler distributions, T4 and T5.

Tables 2 and 3 data might suggest little advantage to approximations more elaborate than P4. Even though σ and $\sigma_{1.6}$ are already small for P4(B), [3] ~ 4 and ~ 3 kJ mol^{-1}, respectively, and the more complex QP4B reduces it by only $\sim 30\%$, closer study indicates this not to be the case. The major

[3] This notation indicates Case B of the P4 model.

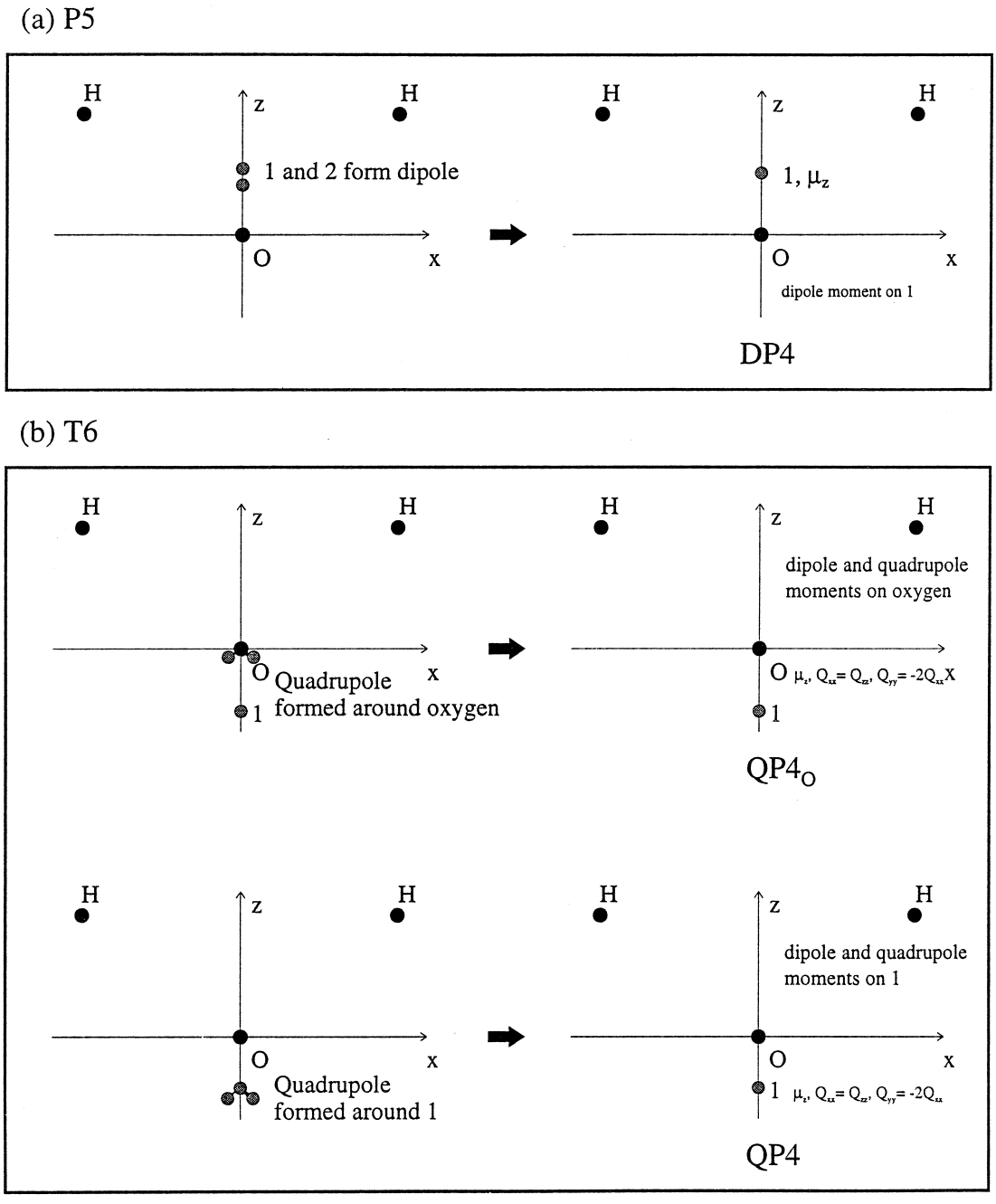

Fig. 2. Schematic diagrams illustrating the optimal P5 and T6 models, explicitly exhibiting their limiting characteristics: (a) in P5 the two shell charges coalesce to form the dipolar DP4 distribution; (b) in T6, the 'lone pair' shell charges coalesce with either the O-atom charge or the other shell charge forming the $QP4_O$ or QP4 distributions respectively. In $SQP4_O$ (not illustrated), the shielded distributed charge is located in the region $z_1 > 0$.

contributions to σ are from the spheres nearer to O, $R \leq 2.5$ Å. Table 4 presents the R dependence of three measures of the error in V, $\sigma(R)$, $\Delta V_{\mathrm{LUB}}(R)$ and $\Delta V_{\mathrm{GLB}}(R)$; $\sigma(R)$ is the rms deviation on the

Table 4
Different measures of the goodness of fit energies in (kJ mol^{-1}) of the various DPCMs considered (see text)

R (Å)	$\sigma(R)$			$\Delta V_{LUB}(R)$				$\Delta V_{GLB}(R)$				$\Delta\rho$ (%)	$\Delta\theta$ (°)
	2.0	2.25	2.5	2.0	2.25	2.5	2.75	2.0	2.25	2.5	2.75	3.0	
Neutrality constraint													
P3	14.8	11.4	9.0	26.0	21.1	16.6	13.0	−32.4	−27.4	−22.9	−16.2	11.3	9.8
P4	7.5	4.1	2.6	3.2	3.5	3.3	2.9	−19.1	−12.1	−8.2	−5.9	2.9	2.6
T4	6.8	3.4	2.1	2.2	2.8	2.8	2.5	−14.6	−9.1	−6.1	−4.3	2.5	2.2
DP4	7.4	3.9	2.3	4.2	4.3	3.6	2.8	−16.6	−10.4	−6.9	−4.9	2.7	2.6
T5	6.7	3.3	2.0	2.1	2.8	2.5	2.2	−14.0	−8.6	−5.7	−4.0	2.3	2.1
P6	5.3	2.1	1.1	−2.4	0.3	1.5	1.5	−9.7	−4.3	−2.1	−1.2	1.2	1.0
QP4	5.4	2.1	1.1	−1.8	−0.0	1.4	1.5	−9.4	−4.1	−2.5	−1.7	1.2	1.0
$QP4_O$	5.7	2.2	1.2	−1.7	0.3	1.3	1.3	−11.3	−5.3	−2.7	−1.8	1.3	1.1
Multipolar constraints													
P3	14.0	10.7	8.9	26.8	21.6	16.9	13.3	−31.7	−28.4	−25.4	−18.1	12.8	9.8
P4	8.4	5.3	3.1	3.2	3.4	5.0	4.6	−19.2	−10.1	−6.1	−4.2	4.9	3.1
T4	7.8	4.5	2.5	2.3	2.9	3.7	3.4	−17.6	−9.2	−5.3	−3.6	3.7	2.7
DP4	8.4	4.4	2.5	3.7	3.9	3.3	3.4	−22.8	−12.3	−7.3	−4.6	3.2	3.0
T5	7.8	4.0	2.2	2.4	3.1	2.7	2.4	−20.2	−10.7	−6.2	−3.9	2.6	2.6
P6	5.5	2.2	0.9	0.1	−0.3	0.3	0.4	−11.1	−5.2	−2.7	−1.5	0.8	0.9
QP4	5.4	2.1	1.1	−1.8	−0.0	1.4	1.5	−9.4	−4.1	−2.5	−1.7	1.0	0.7
$QP4_O$	5.7	2.2	1.2	−1.7	0.3	1.3	1.3	−11.3	−5.3	−2.7	−1.8	0.9	1.0
SP4	5.4	3.6	2.7	7.5	5.4	6.7	5.3	−11.0	−7.4	−5.2	−3.7	5.5	2.6
$SQP4_O$	1.6	1.1	0.8	3.3	3.1	1.9	1.2	−3.7	−2.0	−1.5	−1.1	1.3	0.9
ST6	3.0	2.0	1.4	6.0	4.2	2.9	2.0	−5.2	−3.0	−2.1	−1.6	2.5	1.2

sphere R, $\Delta V_{LUB}(R)$ and $\Delta V_{GLB}(R)$ are the upper and lower bounds, respectively to the quantity $(V^{cal} - V^0)$ on the sphere at R (in all cases sites closer than 1.6 Å to either H-atom are ignored as physically inaccessible [37]). As results for P6, QP4 and $QP4_O$ are almost error free for $R \geq 3.0$ Å (both $\Delta V_{LUB}(R)$ and $\Delta V_{GLB}(R)$ at 3 Å are ~ 1 kJ mol^{-1}), only data for inner spheres are presented.

In relation with the previous work [21], Table 4 contrasts electric fields on the 3-Å sphere for the DPCM models with those determined from the K93 wave function. For QP4(B) and $QP4_O$(B) the rms error in field strength ($\Delta\rho$) and orientation ($\Delta\theta$) are $< 1.0\%$ and $< 1.0°$, respectively. These values for six parameter models, compare favorably with three point multipolar field based distributions (eight parameter) but are inferior to five point (12 parameter) models [21]. [4] Just as for V^{cal}, the two QP4(B) models describe the electric field substantially better than the other models. Both QP4(B) and $QP4_O$(B) improve on the simpler P4 treatment, and for physically important configurations they are better than P6, QP4(A) or $QP4_O$(A).

To further test the QP4(B) models, both were compared with V_0 determined from the PS wave function. Except on the innermost sphere model potentials, derived from K93 data, differ insignificantly from the PS potential; the maximum difference is < 2.9 kJ mol^{-1}, i.e., the models agree with PS data as closely as the K93 and PS calculations agree with one another.

The PS data were used to directly determine alternate parameter sets. Optimal P5 and T6 fits again generated the multipolar representations DP4, QP4 and $QP4_O$ with only slight changes in the parameters.

3.2. Shielded distributed charges

A limitation to all the parameterizations is that $\sigma(R)$ is large on the innermost sphere, 2.0 Å from

[4] These authors found average differences in the magnitude of the field on the 3.0 Å sphere of $-0.2 \pm 1.3\%$; in our model the comparable quantity is much smaller, $-0.03 \pm 1.0\%$.

the O-atom. Since these configurations can be sampled in forming the hydrogen bonded water dimer [37], the associated errors would affect the parameterization of the intermolecular terms in a water force field and seriously limit its reliability in simulating behavior for environments markedly different from those used in establishing the parameters. In addition, the upper and lower bounds to the errors in the potential on the inner spheres are not symmetric, ΔV_{LUB} is systematically less than ΔV_{GLB} in magnitude; in fact, on the innermost sphere the deviations are occasionally **all** negative. Consequently, on the innermost spheres the mean (arithmetic) difference, $\overline{\Delta V(R)}$, between quantum and model electrical energies, $(V_j^0 - V_j^{\mathrm{cal}})$, is negative and differs noticeably from zero whenever $R \leq 2.5$ Å. This systematic bias directly reflects the fact that in these regions water's electronic charge density has not decayed to zero and a probe charge this close to an O-atom senses a net positive charge. It is thus, more realistic to describe the distributed charges as diffuse. Modeling them as exponentially decaying distributions with decay length $1/\gamma_n$, centered at $\mathbf{r}_n$, their individual contributions to the electric potential at $\mathbf{r}$ can be effectively approximated as

$$\frac{q_n}{4\pi\varepsilon_0 D}\left(1 - \mathrm{e}^{-\gamma_n D}\right), \quad D = |\mathbf{r} - \mathbf{r}_n|. \tag{4}$$

Table 5 presents shielded DPCM parameterizations for Case B (neutrality and multipole moment constraints) for P4, T6 and $\mathrm{QP4_O}$ models (shielded models are denoted as SP4, ST6 and $\mathrm{SQP4_O}$, respectively). The fitting errors σ, s and $\sigma_{1.6}$ decrease dramatically, for $\mathrm{SQP4_O}$ by a factor of three. The higher multipole moments are in fair accord with their quantum values. The distributed charges are all negative and located in physically reasonable domains. The final three entries in Table 4 describe the R dependence of the goodness of fit for these shielded model distributions on the inner spheres, limiting consideration to points more than 1.6 Å from either H-atom. In all cases positive and negative deviations are equally likely; the quantity $\overline{\Delta V(R)}$ is essentially zero for all R. For the best model, $\mathrm{SQP4_O}$, the rms deviation is well below 1 kJ mol^{-1}; for no physically accessible configuration do errors exceed 4.0 kJ mol^{-1}. The electric field is well reproduced everywhere, even close to the O-atom. Again excluding points further than 1.6 Å from either H-atom, even on the innermost shell the mean errors in field strength and orientation are only 4% and 2°, respectively, with a maximum orientational deviation of 4.5°. The parameterization is nearly exact at all chemically significant locations and considerably superior to the SP4 model. Water models that ignore shielding must misrepresent the molecule's electric potential at distances important for hydrogen bonding, thus biasing parameterization of intermolecular terms in the water force field, one possible reason for the difficulties in engineering such force fields [23].

Table 5
Shielded (see Eq. (4)) DPCM parameters for Case B with fitting function Eq. (1) minimized subject to both neutrality and multipole moment constraints

	SP4	$\mathrm{SQP4_O}$	ST6
q_{O}	0.8361	3.5893	1.5669
μ_{O}		2.4618	
$\Theta_{yy,\mathrm{O}}$		−0.6596	
q_{H}	0.6015	0.4816	0.5194
q_1	−2.0391	−4.5525	−2.2141
z_1	0.1524	0.15	0.15
γ_1	2.7211	3.1746	2.8345
q_2			−0.1958
z_2			−0.30
y_2			0.49
γ_2			5.2910
σ	2.82	0.89	1.40
s	4.95	1.55	2.47
$\sigma_{1.6}$	2.29	0.68	1.23
Ω_{zzz}	−35.2	−11.7	−28.7
Ω_{xxz}	34.5	8.6	12.5
Φ_{zzzz}	−20.2	3.4	0.1
Φ_{xxzz}	20.3	−3.6	4.3
Φ_{xxxx}	−53.5	−23.2	−23.2

The conventions of Table 2 apply.
The γ_n are reciprocal decay lengths (in $\mathrm{Å}^{-1}$).

The best ST6 model is the $\mathrm{SQP4_O}$ limit. However, unlike the unshielded case, there is also a two shell local minimum. In this case the 'bonding' and 'lone pair' electrons distribute themselves in accord with chemical intuition; with a 'lone pair' angle of $\sim 117°$. It should also be noted that, even if shielding is treated, only P4 and T6 parameterizations lead to model charge distributions that are chemically intuitive. Incorporating shielding into the T4, T5, P5 or P6 pictures is no significant improvement over SP4; the model charge distributions still exhibit

non-intuitive features, i.e., inverted 'lone pairs', positive shell charges, etc. It would appear that, unless the underlying model is truly representative of water's charge distribution, the parameterization is aesthetically unsatisfactory. For genuinely accurate reproduction of water's electric potential, it is crucial both to account for the 'bonding electrons' and to explicitly treat the x–y 'lone pair' asymmetry, either using an ST6 or $SQP4_O$ parameterization.

3.3. Observations

For applications using unshielded DPCMs P5 (DP4) is little improvement over P4 while T6 (QP4 or $QP4_O$) is definitely better. Consequently, in modifying a DPCM like TIP4P [3] to include a point polarizability [38], it would be worthwhile to describe either the distributed site or the O site as a multipole, in effect melding a Barnes-like model [33] with the TIP4P picture. Our analysis provides no clue as to the optimal location of the polarizability site for an unshielded QP4 or $QP4_O$ model since the distributed site in either case is in the 'lone pair' region. An unaltered P3 is an inappropriate basis for polarizable treatment; the errors in such model's electrical potential are substantial even as far as 3.5 Å from O and they are totally unacceptable at distances < 3.0 Å, regions influential for ionic binding. Shielding dramatically improves the utility of simple charge models, yielding representations of an individual molecule's electrical potential that are reliable even at atom-probe distances important for hydrogen bonding. The best of these is $SQP4_O$, where the distributed charge is in the 'bonding' region; for computational reasons it is best suited to development of a rigid water force field. Nearly equivalent results are found with a chemically intuitive ST6 parameterization, which can be used as the basis for a flexible water force field. Parameterization of the intermolecular intershell terms will be based on water multimer properties and tested against the second virial coefficient, ice energetics and liquid water properties.

4. Summary

We have critically evaluated various distributed charge models for water's electric potential with the purpose of assessing whether any is suitable for shell model applications. We have found three novel four site parameterizations that are significant improvements over pure point charge descriptions. Each of these is an optimization of a more complex picture, based on five and six site models which limit naturally to one formed by three point charges and a multipole. In two cases, the multipole is sited at the distributed charge (DP4 and QP4); in the third case it is at the O-atom ($QP4_O$). The quadrupolar models QP4 and $QP4_O$, while only slightly more complex computationally, provide substantially superior fits to water's electrical properties than does the DP4 model. Incorporation of shielding (the diffuse charge distribution) dramatically improves some models. Only shielded models that specifically represent both 'bonding' and 'lone pair' electrons yield aesthetically pleasing parameterizations.

Acknowledgements

This work was supported by a grant from the National Institutes of Health, GM-28643. One of us (JM) would like to thank the American Chemical Society for underwriting a visit to Brandeis University. In addition to providing us her basis set we thank Kersti Hermansson for helpful suggestions.

References

[1] F.H. Stillinger, A. Rahman, Journal of Chemical Physics 60 (1974) 1545.
[2] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular Forces, Reidel, Dordrecht, 1981, pp. 331–342.
[3] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Journal of Chemical Physics 79 (1983) 926.
[4] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, Journal of Computational Chemistry 4 (1983) 187.
[5] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., P. Weiner, Journal of the American Chemical Society 106 (1984) 765.
[6] P.C. Jordan, Biophysical Journal 57 (1990) 1133.
[7] K.A. Duca, P.C. Jordan, Biophysical Chemistry 65 (1997) 123.
[8] J. Aqvist, A. Warshel, Biophysical Journal 56 (1989) 171.
[9] B. Roux, M. Karplus, Biophysical Journal 59 (1991) 961.

[10] B. Roux, B. Prod'hom, M. Karplus, Biophysical Journal 68 (1995) 876.
[11] K.A. Duca, P.C. Jordan, Journal of Physical Chemistry B 102 (1998) .
[12] H.J.C. Berendsen, J. Mavri, Simulating proton transfer processes: quantum dynamics embedded in a classical environment, in: D. Hadži (Ed.), Theoretical Treatment of Hydrogen Bonding, Wiley, Chichester, 1997.
[13] J. Mavri, H.J.C. Berendsen, Journal of Physical Chemistry 99 (1995) 12711.
[14] S.R. Cox, D.E. Williams, Journal of Computational Chemistry 3 (1981) 304.
[15] E. Sigfridsson, U. Ryde, Journal of Computational Chemistry 19 (1998) 377.
[16] T.R. Stouch, D.E. Williams, Journal of Computational Chemistry 14 (1993) 858.
[17] M.M. Francl, C. Carey, L.E. Chirlian, D.M. Gange, Journal of Computational Chemistry 17 (1996) 367.
[18] A.J. Stone, M. Alderton, Molecular Physics 56 (1985) 1047.
[19] A. Wallqvist, P. Ahlström, G. Karlström, Journal of Physical Chemistry 94 (1990) 1649.
[20] C. Millot, A.J. Stone, Molecular Physics 77 (1992) 439.
[21] F. Colonna, J.G. Ángyan, O. Tapia, Chemical Physics Letters 172 (1990) 55.
[22] A.A. Chiavlo, P.T. Cummings, Journal of Chemical Physics 105 (1996) 8274.
[23] A.A. Chiavlo, P.T. Cummings, Journal of Chemical Physics 105 (1996) 8274.
[24] Y.-P. Liu, K. Kim, B.J. Berne, R.A. Friesner, S.W. Rick, Journal of Chemical Physics 108 (1998) 4739.
[25] K. Hermansson, Journal of Chemical Physics 95 (1991) 3578.
[26] K. Hermansson, Journal of Chemical Physics 99 (1993) 861.
[27] H. Partridge, D.W. Schwenke, Journal of Chemical Physics 106 (1997) 4618.
[28] B.G. Dick Jr., A.W. Overhauser, Physical Review 112 (1958) 90.
[29] P.C. Jordan, P.J. van Maaren, J. Mavri, D. van der Spoel, H.J.C. Berendsen, Journal of Chemical Physics 103 (1995) 2272.
[30] N.H. de Leeuw, S.C. Parker, Physical Review B 58 (1998) 13901.
[31] J.G. Ángyán, C. Chipot, International Journal of Quantum Chemistry 52 (1994) 17.
[32] U.C. Singh, P.A. Kollman, Journal of Computational Chemistry 5 (1984) 129.
[33] P. Barnes, J.C. Finney, J.D. Nicholas, J.E. Quinn, Nature 282 (1979) 459.
[34] J.J. Moré, The Lavenberg-Marquardt algorithm: implementation and theory, in: G.A. Watson (Ed.), Numerical Analysis, Springer, Berlin, 1977, pp. 105–116.
[35] R. Fletcher, Practical Methods of Optimization, Vol. 1, Wiley, New York, 1980.
[36] R.J. Woods, M. Khalil, W. Pell, S.H. Moffat, V.H. Smith Jr., Journal of Computational Chemistry 11 (1990) 297.
[37] G.A. Jeffrey, An Introduction to Hydrogen Bonding, Oxford Univ. Press, Oxford, UK, 1997.
[38] M. Sprik, M.L. Klein, Journal of Chemical Physics 89 (1988) 7556.
[39] S.A. Clough, Y. Beers, G.P. Klein, L.S. Rothman, Journal of Chemical Physics 59 (1973) 2254.
[40] J. Verhoeven, A. Dymanus, Journal of Chemical Physics 52 (1970) 3222.