J|A|C|S
ARTICLES
Published on Web 12/12/2003

Ab Initio Molecular Dynamics Study of Formate Ion Hydration

Kevin Leung* and Susan B. Rempe

Contribution from the Sandia National Laboratories, MS 1415 and 0316, Albuquerque, New Mexico 87185

Received May 21, 2003; E-mail: kleung@sandia.gov

Abstract: We perform ab initio molecular dynamics simulations of the aqueous formate ion. The mean number of water molecules in the first solvation shell, or the hydration number, of each formate oxygen is found to be consistent with recent experiments. Our ab initio pair correlation functions, however, differ significantly from many classical force field results and hybrid quantum mechanics/molecular mechanics predictions. They yield roughly one less hydrogen bond between each formate oxygen and water than force field or hybrid methods predict. Both the BLYP and PW91 exchange correlation functionals give qualitatively similar results. The time dependence of the hydration numbers are examined, and Wannier function techniques are used to analyze electronic configurations along the molecular dynamics trajectory.

1. Introduction

The carboxylate ($RCOO^-$) functional group is ubiqitous in biology; all amino acids contain at least one RCOOH acid group. Predicting the thermodynamics and kinetics of acid−base behavior of biological species in aqueous solutions remains an active area of research.[1] $RCOO^-$ also constitutes the polar headgroup in many surfactant molecules, and its interaction with water is important for the formation of lamella phases, vesicles, and micelles.[2,3] Lubrication and biomimetic applications of RCOOH group coated surfaces and RCOOH terminated self-assembled monolayers are also of increasing technological interest.[4]

Despite its importance, there has been limited experimental work on the hydration structure of the formate ion ($HCOO^-$), which is the simplest species containing $RCOO^-$. Nuclear magnetic resonance experiments on carboxylate acids suggest a hydration number of 5−6.5 per carboxylate group.[5] Recently, X-ray and neutron scattering have been performed on 15 mol % NaHCOO[6] and 40 mol % KHCOO[7] aqueous solutions. Infrared (ATR-IR double difference) measurements have also been performed at a series of KHCOO concentrations.[7] From these experiments, it was inferred that the hydration number is below 2.5 per formate oxygen atom for formate concentrations of 15 mol % and lower. The hydration structure of aqueous $HCOO^-$ at lower dilution has yet to be reported.

As a result, structural information about this aqueous ion has come predominantly from computer simulations using molecular force fields in all or part of the calculations. The properties that have received the most attention are N_w, the number of water molecules in the first hydration shell of each solute carboxylate oxygen (i.e., the hydration or coordination number); $g_{O_s-O_w}(r)$ and $g_{O_s-H_w}(r)$, the pair correlation functions between each solute $HCOO^-$ oxygen (O_s) and the oxygen/hydrogen sites of water molecules (O_w, H_w), respectively; and some orientational distributions.

Molecular force fields are typically parametrized using gas phase cluster energies, solvation free energies, or both. Early work by Jorgensen and Gao[8] used Monte Carlo techniques and "optimized potentials for liquid simulations" (OPLS) type empirical force fields to treat the hydration of the simplest $RCOO^-$ species, namely the formate and acetate ions. The carboxylate ion force fields featured fixed bond lengths and angles, no electronic polarizability, and a united atom treatment of −CH and $-CH_3$ groups. The TIP4P potential[9] described the water molecules. The interaction between water and formate ions had the well-known form

$$H_{\text{int}} = \sum_{\alpha i \beta}\left\{\frac{q_\alpha q_\beta}{R_{\alpha i\beta}} + 4\epsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{R_{\alpha i\beta}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{R_{\alpha i\beta}}\right)^{6}\right]\right\} \qquad (1)$$

where α denotes the atomic sites on $HCOO^-$, β labels the interacting sites on water molecule i, $R_{\alpha i\beta}$ is the distance between the respective $HCOO^-$ and water sites, and q_α and q_β are fractional point charges that mimic the electrostatic charge distributions associated with the interacting sites. The Lennard−Jones[10] diameter and coupling strength between sites α and β, $\sigma_{\alpha\beta}$ and $\epsilon_{\alpha\beta}$, mimic Pauli repulsion and dispersion forces and usually derive from mixing the pure species σ_α, σ_β, ϵ_α, and ϵ_β according to some combinatorial rules. Using this force field approach, sharply peaked correlation functions were found that

(1) Ivanov, I.; Klein, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 13380. Davis, J. E.; Doltsinis, N. L.; Kirby, A. J.; Roussev, C. D.; Sprik, M. *J. Am. Chem. Soc.* **2002**, *124*, 6594.
(2) Shah, D. O. In *Micelles, Microemulsions, and Monolayers: Quarter Century Progress at the University of Florida*; Shad, D. O., Ed.; Marcel Dekker: New York, 1998; pp 1−52.
(3) Ulman, A. *Introduction to Ultrathin Films, from Langmuir−Blodgett Films to Self-Assembly*; Academic Press: Boston, MA, 1991.
(4) Liu, N. G.; Assink, R. A.; Brinker, C. J. *Chem. Commun.* **2003**, *3*, 370.
(5) Kuntz, I. D. *J. Am. Chem. Soc.* **1971**, *93*, 514; see also ref 8.
(6) Kameda, Y.; Mori, T.; Nishiyama, T.; Usuki, T.; Uemura, O. *Bull. Chem. Soc. Jpn.* **1996**, *69*, 1495.
(7) Kameda, Y.; Fukuhara, K.; Mochiduki, K.; Naganuma, H.; Usuki, T.; Uemura, O. *J. Non-Cryst. Solids* **2002**, *312−314*, 433.
(8) Jorgensen, W. L.; Gao, J.-L. *J. Phys. Chem.* **1986**, *90*, 2174.
(9) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(10) Lennard-Jones, J. E. *Trans. Faraday Soc.* **1932**, *28*, 334.

10.1021/ja036267q CCC: $27.50

yielded a hydration number of 3.6 for the formate ion and 3.4 for acetate. The aqueous acetate ion was shown to exhibit correlation functions very similar to aqueous $HCOO^-$.[8]

Other carboxylate ion force fields have functional forms similar to that of ref 8; the main differences are the numerical values of the formate parameters q_O, ϵ_O, and σ_O. Thus, another early seminal work by Alagona, Ghio, and Kollman,[11] which also used the TIP4P model for water, focused on the aqueous acetate ion and reported similar structural properties and a slightly lower hydration number of ~3. A higher N_w of 3.4 is observed if q_O, ϵ_O, and σ_O used in this acetate simulation are directly transferred to the formate ion.[12] This conforms to the expectation that steric hindrance arising from the methyl group on acetate ion should yield a lower hydration number than the formate ion. Similar correlation functions are obtained with other fixed charge water models such as SPC/E,[13] which predict an acetate hydration number of 3.3.[14]

While the water models discussed above neglect electronic polarizability, water is known to be highly polarizable; its gas-phase dipole moment is up to 40% smaller than that in the liquid phase.[15] Polarizable water force fields, which have electrostatic dipole moments that can adapt to instantaneous water configurations, have also been considered.[14,16] Using one such force field, POL3,[17] the acetate hydration number dropped slightly, to 3.1.[14] Earlier use of polarizable force fields also suggested some reduction of $RCOO^-$ hydration numbers in amino acid ions.[16,18] Nevertheless, the hydration structure of $RCOO^-$ predicted using nonpolarizable force fields for both water and $RCOO^-$ has been widely accepted, and these force fields are used in numerous simulations of biological[19−21] and other materials containing carboxylate functional groups.[22]

An alternative to the empirical force field approach is to use ab initio methods that do not require the parametrization of water−water and water−solute interactions. Due to the expense of purely ab initio molecular dynamics,[23] where the valence electrons of all atomic species are treated quantum mechanically, only the hydration structures of a few species have been studied using trajectories generated entirely with ab initio forces. In this light, the hybrid ab initio quantum mechanics/molecular mechanics (QM/MM)[24] approach offers a promising alternative. It treats the localized solute region accurately, with quantum mechanical methods, while outlying solvating molecules are modeled using efficient empirical force fields.

Ab initio QM/MM approaches generally apply either Hartree−Fock (HF) or density functional theory (DFT) to treat the solute. In virtually all ab initio QM/MM hydration studies, solvent water molecules are modeled using nonpolarizable empirical force fields. The interactions between the solvent and solute are analogous to eq 1 but with the ab initio electron density replacing the solute partial charges. A set of $\{\epsilon,\sigma\}$ for each *solute* atom otherwise being treated quantum mechanically is thus required.[24] The solvent−solute pairwise radial distribution functions, $g(r)$'s, are sensitive to these σ's and ϵ's as well as to the solvent Lennard−Jones parameters. To obtain better agreement with experiments and/or quantum chemistry calculations in terms of solvation free energies and gas-phase cluster energies, the σ's for solute atoms are often increased ~10% or more relative to those in empirical force field treatments.[25] For example, Cubero et al.[26] used HF to treat the acetate ion, with larger σ's than given in the OPLS force fields, and found that the first hydration shell is pushed farther away from the acetate ion than OPLS predicted. Another recent QM/MM calculation,[12] based on a plane-wave basis and PW91 exchange correlation[27] and applied to the formate ion, adopts $\{\epsilon,\sigma\}$ similar to those of the OPLS potentials and obtains $g(r)$'s much closer to the OPLS and other force field predictions.[8,11] Compared to the carboxylate ion treatments in ref 8, ab initio QM/MM allows the solute bond lengths and angles to fluctuate; no united atom approximation is used, and since the solution to the Kohn Sham or Hartree−Fock equations adiabatically follows the nuclear configuration, zero frequency electronic polarizability is always present. Nevertheless, the hydration numbers of both these QM/MM calculations lie in the range 3.4−3.6 per formate oxygen, similar to refs 8 and 11.

In summary, empirical force fields and the prevalent QM/MM methods predict a wide range of peak positions (2.65−2.95) and peak heights (2.3 to 3.4) in the aqueous formate ion $g_{O_s-O_w}(r)$. The predicted hydration number of 3.4−3.6 is, however, robust for all cases. Each formate ion oxygen atom is predicted to form roughly *one more* hydrogen bond with water molecules than is reported in recent experiments.[6,7]

This discrepancy between experiments and theory may have significant implications when predicting $H_2O-RCOO^-$ group interactions important in acid−base chemistry such as those in glycine, the simplest amino acid containing the $RCOO^-$ functional group. QM/MM calculations have suggested that the intramolecular proton-transfer tautomerization of neutral form → zwitterion aqueous glycine has a free energy barrier of order 1 kcal/mol.[28,29] The experimental value is ~7 kcal/mol.[30] QM/MM methods also estimate glycine zwitterion hydration free

(11) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1986**, *108*, 185.
(12) Leung, K. (unpublished).
(13) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(14) Meng, E. C.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 11460.
(15) Silvestrelli, P. L.; Parrinello, M. *Phys. Rev. Lett.* **1999**, *82*, 3308.
(16) Stamato, F. M. L. G.; Goodfellow, J. M. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1986**, *13*, 277 and references therein.
(17) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 6208.
(18) Recent work on force fields with fluctuating charges also appears promising. For example, see: Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R. H.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515.
(19) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607.
(20) Alagona, G.; Ghio, C.; Kollman, P. A. *THEOCHEM* **1988**, *166*, 385.
(21) Computer simulations of aquous amino acid zwitterions and related species that use force fields similar to ref 8 include: Smith, P. E.; Dang, L. X.; Pettitt, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 67. Gao, J. *J. Am. Chem. Soc.* **1995**, *117*, 8600. Nagy, P. I.; Takács-Novák, K. *J. Am. Chem. Soc.* **1997**, *119*, 4999. Price, D. J.; Roberts, J. D.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1998**, *120*, 9672. Marlow, G. E.; Pettitt, B. M. *Biopolymers* **2003**, *68*, 192.
(22) Shelley, J. C.; Sprik, M.; Klein, M. L. *Langmuir* **1993**, *9*, 916. da Rocha, S. R. P.; Johnston, K. P.; Rossky, P. J. *J. Phys. Chem. B* **2002**, *106*, 13250. Koike, A.; Yoneya, M. *J. Chem. Phys.* **1996**, *105*, 6060.
(23) For a recent overview, see: Sprik, M. *J. Phys. Condens. Matter* **2000**, *12*, A161.
(24) (a) Warshel and Levitt, J. *Mol. Biol.* **1976**, *103*, 227. (b) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(25) Gao, J. L.; Freindorf, M. *J. Phys. Chem. A* **1997**, *101*, 3182.
(26) Cubero, E.; Luque, F. J.; Orozco, M.; Gao, J.-L. *J. Phys. Chem. B* **2003**, *107*, 1664.
(27) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(28) Tuñón, I.; Silla, E.; Millot, C.; Martins-Costa, M. T. C.; Ruiz-López, M. F. *J. Phys. Chem. A* **1998**, *102*, 8673.
(29) See also Tuñón, I.; Silla, E.; Ruiz-López, M. F. *Chem. Phys. Lett.* **2000**, *321*, 433.
(30) This is deduced from the 14.3 kcal/mol activation barrier associated with converting the zwitterion to the neutral form, and the 7.7 kcal/mol free energy difference between them. See: Slifkin, M. A.; Ali, S. M. *J. Mol. Liq.* **1984**, *28*, 215. Haberfield, P. *J. Chem. Educ.* **1980**, *57*, 346.

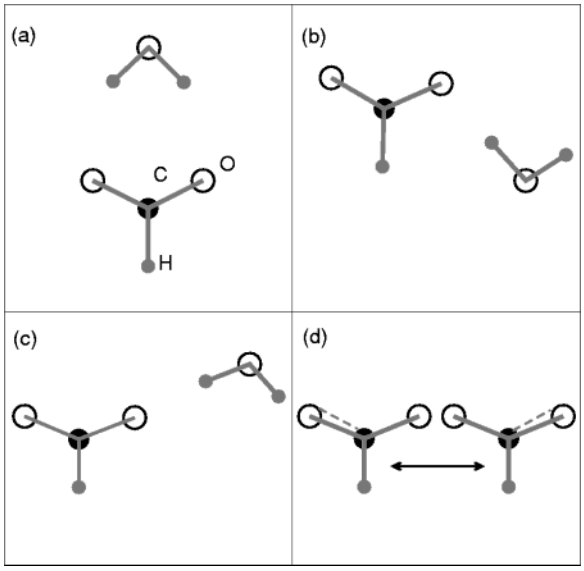

Figure 1. (Panels a−c) Cyclic, anti, and syn $H_2O-HCOO^-$ clusters. (Panel d) $HCOO^-$ resonance structures that may be stabilized due to fluctuations in the water polar environments.

energies that are much too large.[31] While many factors may contribute to these discrepancies, the inability of current QM/MM simulations to produce the experimental $HCOO^-$ hydration number cannot be ruled out as another indication of some problem in the methodology.

In view of this, it is highly desirable to have a database of ab initio simulation results of small molecules and ions that contain different functional groups. Such benchmarks exist for NH_4^+ ion,[32] hydroxide,[33] methanol,[34] dimethyl phosphate,[35,36] and metal cations[37−39] in water, among other species; to our knowledge, $RCOO^-$ has not received similar treatment.

The electronic configurations of aqueous $HCOO^-$ are also of intrinsic theoretical interest. In the gas phase, the two oxygen atoms are equivalent, and the net negative charge is delocalized over both of them (Figure 1d). In a polar solvent like water, the instantaneous environments of the two formate oxygens are different due to fluctuations in their respective solvation shells. It is of interest to investigate whether, to what extent, and on what time scale fluctuations in first solvent shells might stabilize excess charge on one of the oxygens[40] and whether the hybridization of the $HCOO^-$ covalent bonds change as a result.

In this work, we address these issues by computing hydration structures of the aqueous $HCOO^-$ ion using ab initio molecular dynamics simulations. Section 2 describes the computational details. $HCOO^-−H_2O$ interactions in the gas phase are considered in Section 3. Section 4 reports aqueous phase ab initio molecular dynamics results and compares them to experiments, and the last section concludes the paper with discussions of possible explanations of the discrepancy between ab initio, molecular force field, and QM/MM results.

2. Computational Details

We employ two different implementations of gradient-corrected exchange correlation functionals to test their internal consistency. Ab initio molecular dynamics simulations of aqueous formate ion are performed using the Car−Parrinello molecular dynamics (CPMD)[41] code and the Vienna ab initio simulation package (VASP).[42] The exchange correlation functionals BLYP[43,44] and PW91[27] are used with these codes, respectively. Both BLYP[45,46] and PW91[47,48] have been shown to give liquid water structure in reasonable agreement with experiments; the energies of water dimers and of an ordered ice structure they predict also have been tabulated.[49] The simulation cells are cubic, of lateral dimensions 9.9339 Å and 11.7889 Å. They contain one $HCOO^-$ ion and 31 and 53 water molecules, respectively. To the extent that counterions can be neglected, they correspond to formate ion concentrations of 1.7 and 1.0 M. Their volumes correspond to a water density of 1.00 g/cm^3 and the experimental $HCOO^-$ partial molar volume of 51 Å^3.[50] Γ-point Brillouin zone sampling is used in all DFT calculations.

In the CPMD simulations, the wave function energy cutoff is set at 70 Ry, the fictitious electron mass is 800 au, and a time step of 5 au (0.12094 fs) is used. VASP simulations utilize 0.4 fs time steps, ultrasoft pseudopotentials,[51] and a 400 eV (29.4 Ry) energy cutoff. VASP performs dynamics by solving the Kohn−Sham electronic densities and wave functions at each time step, and the total energy is converged to 10^{-4} eV (7.35×10^{-6} Ry). The deuterium nuclear mass is used for protons in both water molecules and the formate ion.

Initial configurations for ab initio molecular dynamics simulations are generated using 25 ps classical mechanics molecular dynamics trajectories with the formate force fields of ref 8, the SPC model for water,[52] the canonical (i.e., constant temperature) ensemble, and a 0.5 fs time step. Then ab initio trajectories are run at the target temperature using a Nosé thermostat,[53] and statistics are collected after discarding the first 1 or 2 ps. The exception is one BLYP simulation used to investigate the effect of varying initial conditions. In this trajectory, we further equilibrate the force field generated configuration with a 4 ps VASP simulation using a large (1 fs) time step and constant velocity rescaling before collecting statistics.

BLYP gas-phase cluster calculations are performed with Gaussian98[54] using a 6-31++G** basis set. Basis set superposition errors with this relatively large basis are expected to fall well below the error bars inherent in the method,[55] and thus, no corrections for superposition errors are made. PW91 gas-phase calculations are performed using VASP with cubic simulation cells of lateral sizes up to 22 Å and monopole/dipole corrections to the energies to account for spurious

(31) Shoeib, T.; Ruggiero, G. D.; Siu, K. W. M.; Hopkinson, A. C.; Williams, I. H. *J. Chem. Phys.* **2002**, *117*, 2762.
(32) Brugé, F.; Bernasconi, M.; Parrinello, M. *J. Am. Soc. Chem.* **1999**, *121*, 10883.
(33) Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **1995**, *103*, 150.
(34) van Erp, T. S.; Meijer, E. J. *Chem. Phys. Lett.* **2001**, *333*, 290.
(35) Kuo, I.-F.; Tobias, D. J. *J. Phys. Chem. B* **2001**, *105*, 5827.
(36) Schwegler, E.; Galli, G.; Gygi, F. *Chem. Phys. Lett.* **2001**, *342*, 434.
(37) Rempe, S. B.; Pratt, L. R. *Fluid Phase Equilib.* **2001**, *183−184*, 121.
(38) Rempe, S. B.; Pratt, L. R.; Hummer, G.; Kress, J. D.; Martin, R. L.; Redondo, A. *J. Am. Soc. Chem.* **2000**, *122*, 966. Rempe, S. B.; Asthagiri, D.; Pratt, L. R. *Phys. Chem. Chem. Phys.*, in press. Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B. *J. Am. Chem. Soc.*, in press.
(39) Ramaniah, L. M.; Bernasconi, M.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 1587.
(40) For discussion of the case of I_3^-, see: Zhang, F. S.; Lynden-Bell, R. M. *Phys. Rev. Lett.* **2003**, *90*, 185505.
(41) Hütter, J.; Ballone, P.; Bernasconi, M.; Focher, P.; Fois, E.; Goedecker, S.; Marx, D.; Parrinello, M.; Tuckerman, M. *CPMD*, version 3.4; MPI für Festkörperforschung and IBM Research Laboratory (1990−1998), Copyright IBM 1990, 1997, and Copyright MPI 1997.
(42) Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169; *Comput. Mater. Sci.* **1996**, *6*, 15.
(43) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(44) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(45) Sprik, M.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1996**, *105*, 1142.
(46) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2002**, *116*, 10372.
(47) Rempe, S. B. (unpublished).
(48) Vassilev, P.; Hartnig, C.; Koper, M. T. M.; Frechard, F.; van Santen, R. A. *J. Chem. Phys.* **2001**, *115*, 9815.
(49) Hamann, D. R. *Phys. Rev. B* **1997**, *55*, 10157.
(50) Aue, D. H.; Webb, H. M.; Bowers, M. T. *J. Am. Chem. Soc.* **1976**, *98*, 318.
(51) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892.
(52) Toukan, K.; Rahman, A. *Phys. Rev. B* **1985**, *31*, 2643.
(53) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511; *Mol. Phys.* **1984**, *52*, 255.
(54) Frisch, M. J., et al. *Gaussian 98*, revision A.2; Gaussian Inc.: Pittsburgh, PA, 1998.
(55) Smith, B. J.; Radom, L. *Chem. Phys. Lett.* **1994**, *231*, 345.

Table 1. Gas (g) and Aqueous (aq) Phase Covalent Bond Lengths and Angles of the $HCOO^-$ Ion, Computed/Measured Using Different Methods

	d_{C-H}, Å	d_{C-O}, Å	O–C–O angle, deg
MP2(g)[a]	1.133	1.258	130.4
PW91(g)	1.144	1.264	130.3
BLYP(g)	1.150	1.273	130.4
PW91(aq)	1.118	1.273	125.5
BLYP(aq)	1.117	1.283	125.9
expt(aq)[b]	1.07	1.272	118

[a] Reference 57. [b] X-ray/neutron scattering results for 15 mol % NaHCOO.[6] The uncertainty in the experimental bond lengths is ∼0.01 Å.

Table 2. Energies and Hydrogen Bond Lengths (*b*) of Three $H_2O-HCOO^-$ Clusters, Computed Using Different Methods[a]

	cyclic		anti		syn	
	E (kcal/mol)	*b* (Å)	*E* (kcal/mol)	*b* (Å)	*E* (kcal/mol)	*b* (Å)
BLYP	−18.9	2.031	−16.8	1.656	−16.1	1.731
PW91	−20.1	1.959	−18.1	1.558	−15.7	1.705
MP2[b]	−19.9	2.023	−17.4	1.674	−16.4	1.747

[a] PW91 results are obtained using periodically replicated unit cells of lateral size 22 Å; the convergence with respective to system size is estimated to be ∼0.2 kcal/mol. [b] Reference 57.

image couplings;[56] these $H_2O-HCOO^-$ binding energies are found to converge to within 0.2 kcal/mol. Basis set superposition errors do not occur in implementations based on plane waves, as in the VASP code.

3. Gas Phase Cluster Results

We compare the DFT description of formate−water interactions with an all-electron Hartree−Fock calculation followed by a Møller−Plesset second-order (MP2) correlation energy correction. Table 1 lists the covalent bond lengths and angles of an isolated $HCOO^-$ ion computed using BLYP and PW91 exchange correlation functionals, as well as the corresponding MP2 predictions.[57] The three ab initio methods agree to within 1% of each other.

Table 2 lists the optimal energies and hydrogen bond lengths of three $H_2O-HCOO^-$ clusters (Figure 1a−c) that have been known to form low energy configurations.[58] The hydrogen bond (i.e., $H_2O-HCOO^-$ binding) energies computed using different ab initio formulations agree to within ∼7% of each other. The BLYP hydrogen bond lengths are close to MP2 predictions, while PW91 hydrogen bond lengths differ from those by up to 0.12 Å. The variation in hydrogen bond energies among various ab initio methods is not surprising; even larger fractional discrepancies have been reported for the water dimer binding energy.[49] It is more significant that the energy and bond length differences and ordering are similar for the three clusters, regardless of the method used; this suggests that, in terms of the $H_2O-HCOO^-$ interaction, the two DFT exchange correlation functionals are equally valid. We also note that the classical force fields of ref 8 give similar gas cluster energies. As will be shown, however, these force fields yield aqueous phase $g(r)$'s rather different from ab initio molecular dynamics predictions. The difference is likely due to (a) lack of collective effects in the force fields and (b) their inability to depict accurately the potential energy surface away from the local minima in gas-phase clusters.

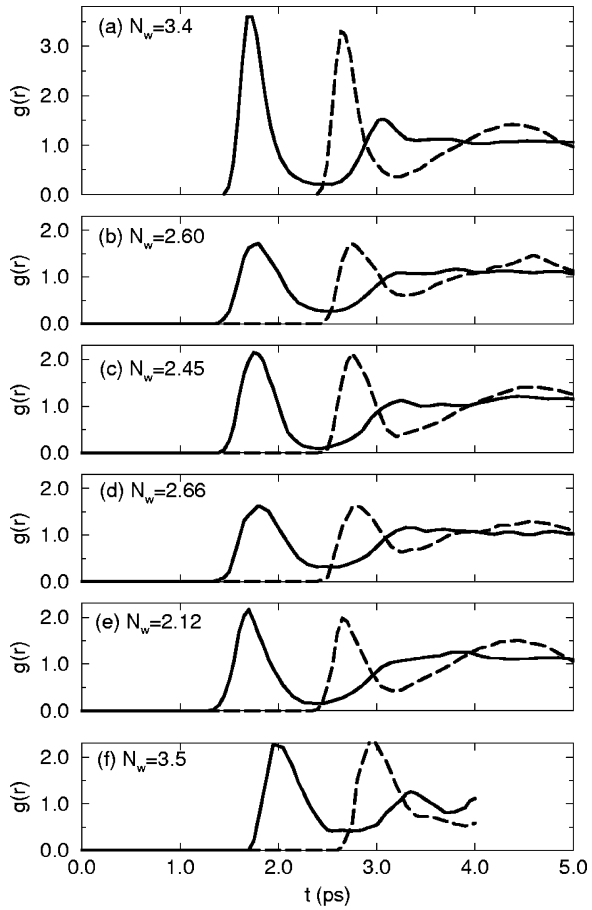

Figure 2. $g_{O_s-H_w}(r)$ (solid line) and $g_{O_s-O_w}(r)$ (dashed line). (a) OPLS, T = 298 K, 45 ps;[8] (b) BLYP, 31 H_2O molecules, T = 300 K, averaged over 14.2 ps; (c) BLYP, 53 H_2O molecules, T = 300 K, 7.8 ps; (d) BLYP, 31 H_2O molecules, T = 350 K, 8.7 ps; (e) PW91, 31 H_2O molecules, T = 350 K, 14.6 ps; (f) QM/MM for acetate ion, T = 298 K, adapted from ref 26.

4. Aqueous Phase Results

We find that PW91 aqueous $HCOO^-$ trajectories at $T \approx 300$ K exhibit slow dynamics; in particular, $HCOO^-$ rotation is effectively frozen in the simulation time scale. As a result, the PW91 trajectory is performed at an elevated temperature of T = 350 K. We note that such sluggish dynamics does not occur for pure water simulations under standard conditions[47,48] or with constant velocity rescaling during equilibration.

The average covalent bond lengths and angles in the aqueous simulations are listed in Table 1. As in the gas phase, BLYP and PW91 results agree to within 1% in all cases. Upon transferring $HCOO^-$ from the gas to the aqueous phase, both BLYP and PW91 predict that the C−O bond length increases, the C−H bond length decreases, and the O−C−O angle decreases. The predicted aqueous phase C−O bond lengths are in good agreement with X-ray and neutron scattering experiments, while the ab initio formate C−H bond length and O−C−O angle are too large by ∼5%. Note that these experiments are performed in concentrated aqueous sodium formate solutions at 15 mol %; in other words, there is one hydrated $Na^+/HCOO^-$ pair for every six or seven water molecules. Therefore, association between Na^+ and $HCOO^-$ cannot be neglected, and they may affect molecular geometries in these experiments.

(56) Makov, G.; Payne, M. C. *Phys. Rev. B* **1995**, *51*, 4014.
(57) Markham, G. D.; Trachtman, M.; Bock, C. L.; Bock, C. W. *THEOCHEM* **1998**, *455*, 239.
(58) Lukovits, I.; Karpfen, A.; Lischka, H.; Schuster, P. *Chem. Phys. Lett.* **1979**, *63*, 151.

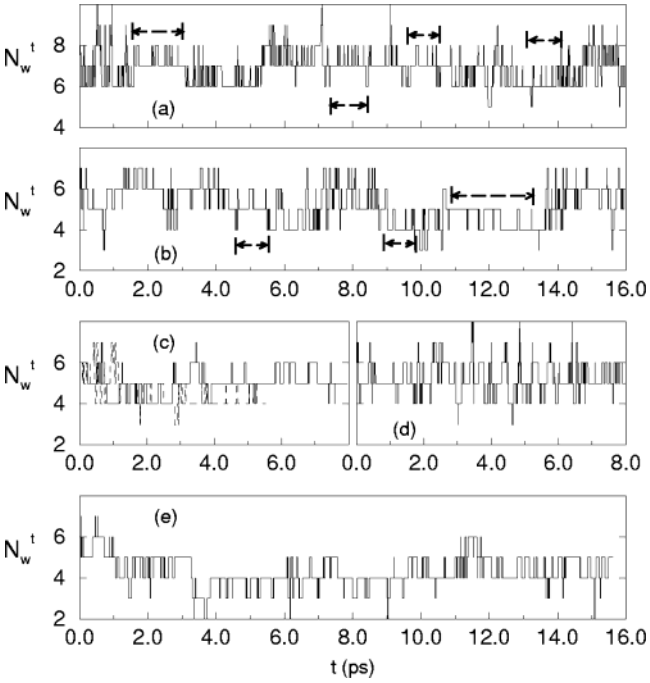

Figure 3. Time dependence of the number of water molecules in the first solvation shell for the two formate oxygens combined (N_w^t). The trajectories correspond to those in Figure 2, except for the dashed line in Figure 3c which describes a trajectory originating from a different initial condition. In panels a and b, a few of the periods where N_w^t spends significant time fluctuating between two values are highlighted. In some cases, only a portion of the trajectory is shown. The first 1 or 2 ps of these trajectories are discarded when computing $g(r)$'s and N_w's.

Figure 2 compares $g_{O_s-O_w}(r)$ and $g_{O_s-H_w}(r)$ computed using ab initio molecular dynamics, QM/MM predictions, and empirical classical force fields. Figure 3 depicts N_w^t, the instantaneous hydration number of both formate oxygens combined, where $\langle N_w^t \rangle = 2N_w$. These figures represent the main findings of this work.

In Figure 2a, we first reproduce the $g_{O_s-H_w}(r)$ and $g_{O_s-O_w}(r)$ predicted by Jorgensen and Gao,[8] using their empirical force fields for $HCOO^-$ solvated in 255 TIP4P water molecules at $T = 298$ K and a water density of 1 g/cm^3. The first peak of the O_s-O_w and O_s-H_w $g(r)$'s occurs at 2.65 Å and 1.72 Å, respectively, with height values that exceed 3.2 times the density of bulk water. A second peak in $g_{O_s-H_w}(r)$ is predicted at 3.1 Å. Integrating to the first minimum in $g_{O_s-H_w}(r)$, we find that there are $N_w = 3.4$ water molecules in the first solvation shell of each formate oxygen atom.[59] Using these force fields, we find that the solute−solvent $g(r)$'s are almost unchanged when the simulation cell size is reduced and that N_w values for cells with 31 and 255 water molecules differ by only ∼3%. Similar conclusions concerning the simulation cell size are reached with the dimethyl phosphate ion.[35] To the extent that ion clustering (and the presence of counterions in the $HCOO^-$ case) can be neglected, this suggests that anion concentration changes do not strongly affect first hydration shell structure. Increasing the temperature to 350 K slightly reduces the peak heights in the correlation functions, but has no discernible effect on N_w.

Figure 3a depicts N_w^t over a portion of the trajectory. The fluctuations in this quantity are of a subpicosecond time scale.

(59) In this work, we always use the $g_{O_s-H_w}(r)$ pair correlation function to determine the hydration number; integrating the $g_{O_s-O_w}(r)$ pair correlation function yields a similar N_w, within ∼0.1 of present results.

Table 3. Hydration Numbers (N_w), First Peak Positions (d_x), and First Peak Heights (h_x) of $g_{O_s-X_w}(r)$, Where X Is Either H or O[a]

	a	b	c	d	e	f
N_w	3.4	2.60	2.45	2.66	2.12	3.5
d_H	1.72	1.77	1.77	1.80	1.70	2.00
h_H	3.6	1.7	2.1	1.6	2.1	2.3
d_O	2.65	2.75	2.75	2.80	2.64	2.95
h_O	3.3	1.7	2.1	1.6	2.0	2.4

[a] The peak heights have statistical uncertainties of 0.15. Columns a−f refer to the six calculations listed in the caption of Figure 2.

It is also apparent that a longer time scale exists; the system resides in configurations where the combined hydration number of both formate oxygens (N_w^t) predominantly fluctuates between 6 and 7 and between 7 and 8 water molecules. The sojourn time in each of these N_w^t regimes ranges from a fraction of a picosecond to 2 ps, after which the system undergoes transitions to regimes associated with other pairs of N_w^t values.

These empirical force field results will be the reference against which the following ab initio molecular dynamics results are compared.

The first $g_{O_s-H_w}(r)$ peak occurs at 1.77 Å in both BLYP simulations conducted at $T = 300$ K, at 1.80 Å for $T = 350$ K, and at 1.70 Å for the PW91 trajectory. $HCOO^-$ is a much weaker base than OH^-, and we do not expect to see spontaneous proton transfer from the H_2O molecules to $HCOO^-$. The $g(r)$'s indeed verify that, along the molecular dynamics trajectories, the distance of closest approach between the formate oxygens and the water protons is ∼1.33 Å in all cases. The variation in the first O_s-O_w peak position is larger than that for O_s-H_w, with PW91 predicting 2.64 Å and BLYP ranging from 2.75 to 2.80 Å. PW91 thus yields smaller O_s-H_w hydrogen bond lengths in both gas and aqueous phases; the BLYP values are closer to experimental estimates.[6,7] Despite the strong hydrogen bond between $HCOO^-$ and H_2O (see Table 2), the first peaks of the O_s-H_w and O_s-O_w $g(r)$'s are rather diffuse; they are much less sharply peaked than those predicted using empirical force fields (Figure 2a). Comparing the two BLYP simulation cell sizes (Figure 2b and c), we find that the 53 H_2O simulation yields sharper $g(r)$ features than the 31 solvent molecule simulation. Increasing the temperature by 50 K (Figure 2d) also leads to more diffuse peaks in the $g(r)$'s. The peak heights have a statistical uncertainty of ∼0.15. These results are summarized in Table 3.

The second peak in $g_{O_s-H_w}(r)$ is barely discernible in all ab initio simulations. We find that, with empirical force fields, this second peak consists almost entirely of the protons on water molecules hydrogen-bonded to the formate oxygens through their other protons. In the ab initio trajectories, there are fewer such nonbonding protons and they are less ordered than in empirical force field configurations.

Figure 3b shows that there are large fluctuations in the hydration number during the ab initio trajectories, similar to the case of aqueous Na^+.[37] The sojourn times where N_w^t mainly oscillates between two values are also qualitatively similar to empirical force fields results. The prevalent value of N_w^t, however, now fluctuates between 4 and 5 and between 5 and 6. Within these sojourn times, water molecules in the first solvation shells are usually tightly bound to the formate oxygens and exhibit hydrogen bond lengths less than 2.1 Å. The

fluctuations mostly consist of either (a) one "loosely bound" H_2O entering a solvation shell, briefly forming a longer, tenuous hydrogen bond, and then leaving or (b) one of the tightly bound water molecules briefly departing and then re-entering the solvation shell. In contrast, in the "transition" regimes, *several* water molecules can be loosely bound to the formate oxygens, and faster, larger magnitude fluctuations in N_w^t occur. These transitions are frequently initated when a water proton is simultaneously within the first hydration shells (that is, within ~2.5 Å) of both formate oxygens. The solvation shells can be asymmetric; when $N_w^t = 4(6)$, it is fairly common to find the two formate oxygens forming one and three (two and four) hydrogen bonds, respectively, as opposed to two (three) apiece. N_w^t values of 3 also occur with some frequency in the PW91 trajectory (Figure 3e).

Given these long sojourn times, we vary both the initial conditions and the temperature to check that N_w has reached its equilibrium value. Thus, the initial configurations of the two trajectories depicted in Figure 3c are generated using VASP and classical force field simulations, respectively. They yield $N_w = 2.45$ and 2.25. The latter is sampled using a short trajectory, making statistical uncertainties larger. Nevertheless, the two simulations are in reasonable agreement. Figure 3c also shows that the larger simulation cell decreases the tendency of the formate ion to form six hydrogen bonds for long periods of time. This cell also exhibits slower N_w^t fluctuations, which are consistent with the sharper features in the $g(r)$'s (Figure 2c). The $T = 350$ K BLYP trajectory (Figure 3d) exhibits faster oscillations between $N_w^t = 4$, 5, and 6, but N_w still averages to 2.66, similar to the $T = 300$ K run, where $N_w = 2.60$. This suggests that PW91 ab initio molecular dynamics at $T = 300$ K would yield results similar to those at $T = 350$ K, although the dynamics will be slower. Finally, we note that the formate oxygens can spontaneously form a total of seven hydrogen bonds in all ab initio trajectories, which is the average empirical force field value. This shows that the ab initio and empirical force field trajectories overlap in the regions of configuration space they sample. These tests amply demonstrate that N_w is adequately equilibrated; the smaller ab initio value compared to empirical force fields is not due to the system being trapped in a metastable state. Figure 3 suggests that relatively long simulation times are necessary. For example, if we had stopped the simulation depicted in Figure 3b after 4 ps, we would have concluded that ab initio and empirical force field results are similar.

The statistical uncertainty in N_w in the Figure 2b trajectory is estimated to be ~0.12 water molecules. Therefore, all N_w's predicted in BLYP simulations should be within two standard deviations of each other. Thus, raising the temperature by 50 K or using a larger simulation cell has only small effects on the average values, similar to the case for empirical force fields. The PW91 N_w (Figure 2e) is 20% smaller than that predicted in the BLYP trajectory (Figure 2d) obtained at the same temperature.

As discussed in the Introduction, the neutron scattering experiments of Kameda et al. suggest that the hydration number per formate oxygen is $N_w = 2.2 \pm 0.1$ in 15 mol % NaHCOO.[6] By analyzing infrared spectroscopy results at a series of KHCOO concentrations, they estimated that $N_w \approx 1.9$ at 15 mol % KHCOO and extrapolated an $N_w \approx 2.3$ per formate oxygen at infinite dilution.[7] These N_w are much smaller than the empirical force field predictions. They are consistent, however, with the N_w predicted by the ab initio trajectories shown in Figure 2, and the infinite dilution value of $N_w = 2.3$ is bracketed by the BLYP and PW91 predictions. Note that, in ref 7, there are no estimates based on infrared measurements reported for structural properties such as hydrogen bond lengths at infinite KHCOO dilution.

Finally, Figure 2f depicts QM/MM $g(r)$'s obtained by Cubero et al.,[26] which used Lennard−Jones diameter σ for quantum mechanical atoms that are ~10% larger than the same OPLS parameters.[25] Superficially, the broader first peaks in the pair correlation functions resemble the ab initio results. They occur, however, at larger distances; the first O_s-H_w peak occurs at ~2 Å instead of the ab initio predictions of 1.70 or 1.77 Å. Integrating the $g(r)$'s numerically, we find $N_w \approx 3.5$, similar to the OPLS results. The second peak in $g_{O_s-H_w}(r)$ is also more distinct than ab initio results. Despite the range of parameters used, QM/MM and empirical force field methods consistently yield $3.0 < N_w < 3.6$ for both the formate and acetate ions. In particular, $N_w \geq 3.4$ in all reported simulations of aqueous $HCOO^-$ that use nonpolarizable force fields, values considerably larger than experimental[6,7] and our ab initio molecular dynamics studies predict.

An interesting comparison can be made between $HCOO^-$ and the aqueous dimethyl phosphate ion $(CH_3)_2PO_2^-$. Both species carry a net negative charge of −1, which is delocalized over two oxygen atoms in the gas phase. An ab initio molecular dynamics simulation of aqueous $(CH_3)_2PO_2^-$ by Kuo and Tobias[35] shows that it exhibits a coordination number of $N_w = 2.5$ around each lone oxygen atom, similar to what we find for formate ion, but in contrast to the 2.75 to 3 coordinated waters found in empirical force field calculations performed by them and earlier by Alagona, Ghio, and Kollman,[60] respectively. While the agreement between the ab initio and classical force field results are closer here, still the force field approach predicts a more structured, more well-defined coordination shell containing more water molecules than the ab initio studies. We note that this ab initio simulation ran for a short length of time (only 4.5 ps); thus their results may depend on initial conditions, and sufficient statistics may not have been gathered. Hints of that possibility come from the reported 6-fold coordination of the sodium counterion by Kuo and Tobias,[35] which is larger than what recent ab initio studies utilizing longer simulation times predict.[37] We anticipate that longer simulations may lower the ab initio coordination number for both the dimethyl phosphate and sodium ions, thus increasing the discrepancy with force field results.

Next, we investigate possible reasons ab initio predictions differ from force fields and QM/MM. Three possibilities might be (a) formate−water hydrogen bonds may have significant covalent character; (b) there may be strong hybridization changes in the aqueous formate ion that are not accounted for in classical force fields, which somehow QM/MM also fails to account for; and (c) water polarizability is neglected in the cited force fields and QM/MM treatments. We will examine possibilities a and b herein.

The orientation of water molecules around the formate oxygens is depicted in Figure 4a. The results are qualitatively

(60) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1985**, *107*, 2229.

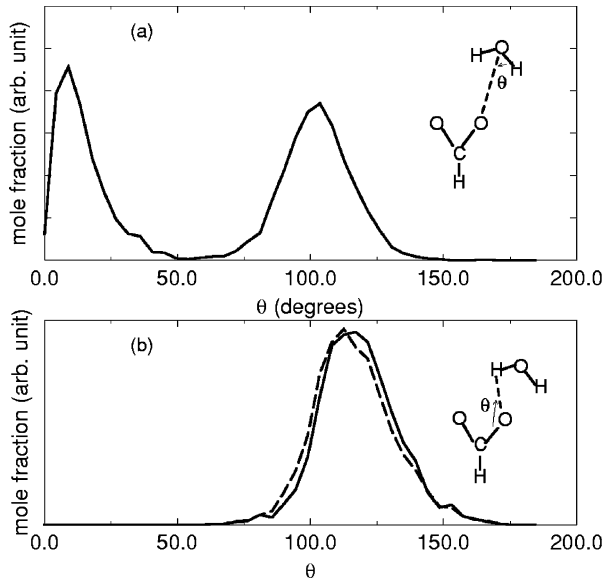

Figure 4. (a) Orientational distribution of water molecules in the first solvation shell of the formate oxygens, computed from the trajectory depicted in Figure 3b. The angle θ is defined in the figure. To be consistent with the literature, contributions from both hydrogens in the water molecule are summed over, which explains the presence of the two peaks at ∼9° and ∼104°. (b) Orientational distribution of the formate ion CO bond with respect to formate−water hydrogen bonds. Only the proton in the water molecule closest to the formate oxygen is counted, and so only one peak exists. Solid line: contribution from formate oxygen hydrogen bonded to two water molecules. Dashed line: contribution from formate oxygens in all configurations.

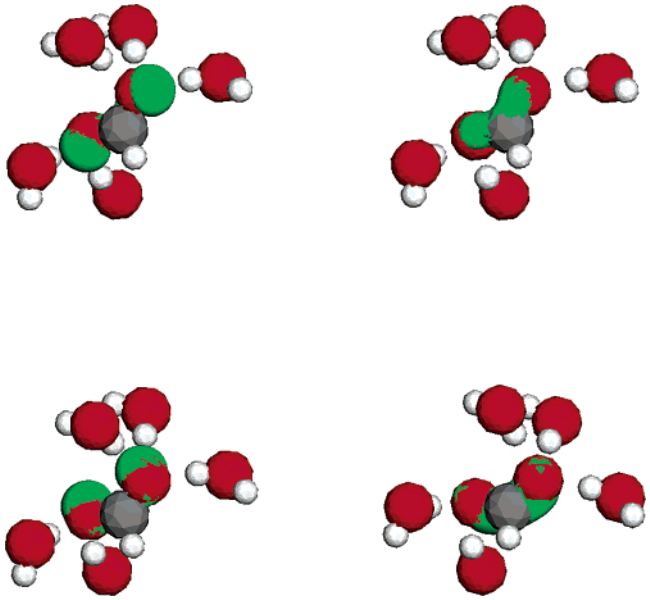

Figure 5. Sample hydration shell water configuration and maximally localized Wannier orbitals along the trajectory described in Figure 2. (Red) O; (gray) H; (dark gray) C; (green) electron density associated with Wannier orbitals. There are three H_2O molecules hydrogen bonded to one of the $HCOO^-$ oxygens and two H_2O molecules bonded to the other. Four Wannier orbitals are associated with each $HCOO^-$ oxygen. Each panel corresponds to the same nuclear configuration and depicts the electron density associated with a Wannier orbital on each oxygen. The orbitals are shown in symmetric pairs to emphasize that the two oxygens exhibit similar electronic configurations.

similar to ab initio molecular dynamics predictions for aqueous dimethyl phosphate[35] and empirical force field results for aqueous acetate.[11] They indicate that one of the two water OH bonds points almost directly at the formate oxygen.

Figure 4b plots another angular distribution, the one formed between the formate C−O bond and the hydrogen bond the formate oxygen makes with water, which sheds some light on the electronic configurations. The distribution peaks at 113° and 115° depending on whether all formate oxygens or only those having two hydrogen bonds are counted. If the solvent fluctuations cause strong charge localization on one of the formate oxygens and the formation of a C−O double bond, and if the hydrogen bonds are strongly covalent, one expects highly directional hydrogen bonds that cause the C−O−H angle to peak at 109.5° and 120° depending on whether the oxygen is sp^3 or sp^2 hybridized, respectively. Figure 4b does not support the existence of such configurations.

In Figures 5 and 6, we confirm the absence of strong charge localization by plotting the maximally localized Wannier functions[61] in two sample configurations. Figure 5 has the formate oxygens forming two and three hydrogen bonds, respectively. Figure 6 has two water molecules in the first solvation shell of each formate oxygen, but a proton in one of the water molecules is only 1.33 Å away from one of the formate oxygens. As has been discussed, such short hydrogen bond lengths occur infrequently. In both cases, four occupied Wannier orbitals are centered around each formate oxygen; therefore, the net ionic charge is clearly localized in the vicinity of the oxygens. In Figure 6, a Wannier orbital overlaps the O and H separated by 1.33 Å, showing that covalent bonding is indeed

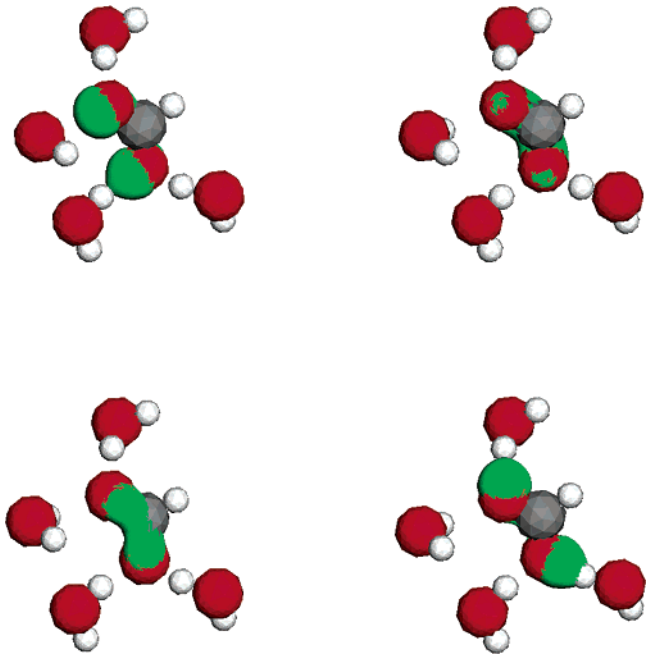

Figure 6. Another sample hydration shell water configuration and maximally localized Wannier orbitals along the trajectory described in Figure 2b. See Figure 5 for key to symbols. There are two H_2O molecules hydrogen bonded to each of the $HCOO^-$ oxygens in this configuration. One of the water protons is at a distance of 1.32 Å from a formate oxygen. Such hydrogen bond lengths occur very infrequently. See text.

present. Despite this, the Wannier orbitals are fairly symmetrically distributed on both formate oxygen atoms; there is no suggestion of C−O double bond formation. A Mulliken charge analysis of this configuration does not indicate substantial charge localization on one of the oxygens. The same is true of

(61) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12847.

Figure 5, where the two formate ions form different numbers of hydrogen bonds. So, the "resonance structure" picture depicted in Figure 1d seems to prevail in the aqueous phase. Despite the discrepancy with empirical force field $g(r)$'s, the asymmetry in the hydration shells is a subtle effect; it may be the result of electrostatic or steric effects, not hybridization changes.

The above discussion suggests that covalent bonding/formate ion hybridization changes are not key factors that influence the ab initio predicted hydration structure. Water polarizability, however, remains a viable reason.

5. Conclusions

The ab initio molecular dynamics simulations of this work show that aqueous formate−water $g(r)$'s exhibit broader peaks and features compared to widely accepted empirical force field predictions. The first O_s-O_w peak is at 2.64−2.80 Å, while $g_{O_s-H_w}(r)$ first peaks at 1.70−1.80 Å.

Using a simulation cell with 53 water molecules and the BLYP exchange correlation functional, a hydration number of 2.45 per formate oxygen is predicted, consistent with recent experiments.[6,7] Varying the temperature and simulation cell size and replacing the BLYP exchange correlation functional with PW91 yields hydration numbers ranging from 2.12 to 2.66. In other words, all ab initio simulations that we have performed predict hydration numbers within 16% of the experimental value. In contrast, widely cited empirical force field results and recent QM/MM simulations yield hydration numbers of ∼3.4. Our simulations and experimental results suggest that empirical force fields and existing QM/MM methods overestimate the number of hydrogen bonds between each formate oxygen and water by about one.

The hydration numbers exhibit picosecond sojourn times in certain subsets of hydration numbers along the molecular dynamics trajectories. Similar dynamical behavior may occur in more complex species that contain the $RCOO^-$ group. Long ab initio trajectories are required to sample the hydration structure adequately in such situations, especially because ab initio simulations typically use initial configurations generated with classical force fields that exhibit significantly different $RCOO^-$ hydration structures.

One way to further validate the DFT simulations in this work is to use the quasi-chemical method[62] to compute the formate hydration free energy and compare it to experiments. This approach also should shed light on the change in hydration free energy as successive water molecules are added to the first hydration shell and thus pinpoint the reason underlying the preference for hydration numbers of 4 and 5 in ab initio molecular dynamics over $N_w = 6$ or 7, which dominates in empirical force field studies. We will pursue this avenue in the future. In the meantime, we have compared different exchange correlation functionals and simulation conditions. The conclusion is that the hydration numbers predicted by ab initio methods all lie outside the range predicted by empirical force fields and hybrid QM/MM methods. It appears that the latter methods[8,11,20−22,26] should be re-examined. Finally, our understanding of carboxylate hydration will greatly benefit from future X-ray and neutron scattering experiments performed at low solute concentrations.

Regardless of how accurate the present-day density functional theory exchange correlation functionals are, to the extent that the QM/MM method can and should be considered an approximation to full ab initio molecular dynamics, the hydration number discrepancy between the two approaches is troubling. This is particularly true because the formate ion is a simple species containing only C, H, and O atoms. While discrepancies between ab initio molecular dynamics and force field (or even QM/MM) predictions of hydration structures have been reported for metal ions[37,38,63] and highly polarizable anions such as bromide,[64] our work shows that even simple, fairly ubiquitous organic ions can exhibit this anomaly. We conjecture that either (a) the functional form of the QM/MM coupling or (b) the water model used may be inadequate. To address the former, more sophisticated electrostatic couplings[65] can be used. Indeed, treating water molecules in the first solvation shell quantum mechanically is known to modify dramatically the pair correlation functions.[66] Regarding water models, applying polarizable water models within the ab initio QM/MM method[12] may be a promising route, because polarizable empirical force fields have been demonstrated to lower the hydration number.

Acknowledgment. We thank Chris Mundy and Marcus Martin for useful suggestions. The Wannier orbital visualizations are performed using the code "dan" developed by Norm Bernstein at Harvard University; we also thank Normand Modine for assistance. This work was supported by the Department of Energy under Contract DE-AC04-94AL85000 and was funded in part by the U.S. Department of Energy's Genomes to Life program (www.doegenomestolife.org) at Sandia. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy's National Nuclear Security Administration.

JA036267Q

(62) Pratt, L. R.; Rempe, S. B. In *Simulation and Theory of Electrostatic Interations in Solution*; Pratt, L. R.; Hummer, G. Eds.; ALP: New York, 1999; pp 172−201.

(63) Tongraar, A.; Rode, B. M. *J. Phys. Chem. A* **2001**, *102*, 10340.

(64) Raugei, S.; Klein, M. L. *J. Am. Chem. Soc.* **2001**, *123*, 9484; *J. Chem. Phys.* **2002**, *116*, 196.

(65) Wesolowski, T.; Muller, R. P.; Warshel, A. *J. Phys. Chem.* **1996**, *100*, 15444.

(66) Tongraar, A.; Rode, B. M. *J. Phys. Chem. A* **2001**, *105*, 506.