J. Phys. Chem. A **2008,** *112,* 12687–12694

Theoretical Prediction of pK_a Values for Methacrylic Acid Oligomers Using Combined Quantum Mechanical and Continuum Solvation Methods

Haitao Dong, Hongbo Du, and Xianghong Qian*

Department of Mechanical Engineering, Campus Delivery 1374, Colorado State University, Fort Collins, Colorado 80523

Received: August 15, 2008; Revised Manuscript Received: October 3, 2008

The acidity constant pK_a for polymeric organic acid is expected to be different from its corresponding monomer value due to the change of chemical environment upon polymerization. Thermodynamic cycles were used to determine the free-energy changes for the proton dissociation processes in aqueous solution and the corresponding pK_a values for monomer methacrylic acid and several similar carboxylic acids. First-principles calculations and continuum solvation model were used to determine the gas-phase and solvation free energies, respectively. A protocol was developed to use the efficient density functional calculations with B3LYP functional instead of the demanding CBS-QB3 method to determine the gas-phase free energies with relative high accuracy, thus enabling the determination of pK_a values for the short oligomers of methacrylic acid. The predicted pK_a values for the dimer and trimer of methacrylic acid are higher by about 0.8 pK_a units than the predicted monomer value.

1. Introduction

The acid dissociation constant, pK_a, is a measure of an acid′s tendency to release a proton in an aqueous solution. This quantity is of great importance in many chemical and biological processes such as ion exchange based separations because it determines the relative concentrations of the acid and its conjugate base at a given pH value. [1] Accurate pK_a prediction remains a major challenge in computational sciences, particularly for polymeric acids with multiple functional groups. In a complex polymeric system involving multiple acidic sites and conformational structures, the pK_a values of the acidic groups are typically different from their monomer counterparts due to the extra bond formed between the monomer units, the additional van der Waals and electrostatic interactions, and any potential hydrogen bonds formed among the functional groups. For example, Forsyth et al. show that two different −COOH groups in the same molecule can have pK_a values of 2 and 9, indicating very different chemical environments.[2] Another experimental study on organosulfonic acid adsorption on mesoporous silica surface shows that pK_a values of the sulfonic groups are dependent on their spatial locations.[3] The sensitivity of pK_a values on structure and environment provides an opportunity to develop efficient acid catalysts by creating optimal catalytic pH values. This pK_a dependence will also affect ion-exchange membranes used for targeted bioseparations.[4] Membrane surfaces are typically modified by attaching, e.g., methacrylic acid polymer nanobrushes using controlled surface modification techniques such as atom-transfer radical polymerization (ATRP).[5] By controlling the length and the density of the polymeric nanolayers, appropriate pore size of the membranes can be obtained for protein or other macromolecule separations. However, the pK_a values of the polymeric nanolayers of weak acids or bases are typically unknown and will be different from their corresponding monomer values. It is therefore critical to determine the pK_a values of the polymeric nanolayers for optimal separation conditions. The present work is an attempt to investigate the pK_a dependence of the short chain polymeric 2-methyl-2-propenoic (methacrylic) acids on their chain length.

Theoretical approaches based on the principles of physical chemistry are used to determine the pK_a values of ionization groups in complex chemical environment. In fact theoretical prediction of pK_a values has long been an active area in computational chemistry.[6−22] However, previous theoretical calculations were carried out for either monomer acids[6] or macromolecules such as proteins.[10,22−24] Typically the former uses quantum mechanical methods whereas the latter solves the classical Possion−Boltzmann equation. In both cases, the solvents are usually treated implicitly using the continuum solvation models. A challenge facing these theoretical approaches is that the errors involved in calculating the free energy ΔG of proton dissociation have to be small since the dissociation constant K_a depends exponentially on ΔG. Further, to date no attempt has been made to predict the pK_a values of polymeric acids, and their dependence as a function of degree of polymerization (DP).

Statistical mechanical methods are usually used for direct calculations of proton dissociation free energies and the corresponding pK_a values.[25] Davies et al. calculated the pK_a value of equatorial P−OH in phosphoranes by evaluating the reversible work needed for proton dissociation. Their predicted value is about 1 pK_a unit higher than the experimental result.[23] The use of thermodynamic cycles, on the other hand, has been applied more broadly for pK_a calculations. This approach indirectly determines free-energy change for proton dissociation in solution via calculating gas-phase free-energy change for proton dissociation and the solvation free energies of the neutral and dissociated species. Thermodynamics cycles were used in this work to determine the pK_a values of methacrylic acid oligomers with DP varying from 1 to 3 because direct quantum mechanical calculation of their free energies in solution is still not possible.

---

* To whom correspondence should be addressed. E-mail: xqian@engr.colostate.edu.

10.1021/jp807315p CCC: $40.75

Two quantum chemical methods, the complete basis set model (CBS) developed by Petersson et al. and Gaussian-*n* model developed by Pople et al., are widely used to calculate gas-phase free energies because of their high degree of accuracy.[26,27] The errors in calculated free energies for a variety of molecules in the gas phase are found to be about 1 kcal/mol. However, these two methods can not be used to calculate the free energies of polymeric methacrylic acids due to the large number of atoms involved and the large basis sets needed. Density functional theory (DFT) based hybrid B3LYP methods were used in our calculations because of their relative high accuracy (1−2 kcal/mol) and capability of treating large systems.[28,29]

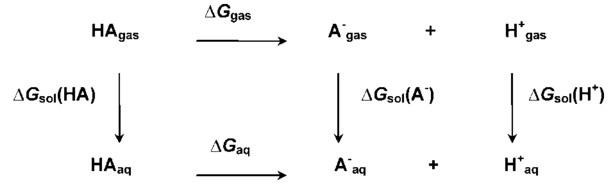
There are two theoretical approaches to obtain solvation free energies for the acids and their conjugate bases. The first method uses explicit solvent molecules to determine the free-energy change associated with the solvation process.[30,31] The second method treats the solvent as a structureless continuum medium with an appropriate dielectric constant.[14,32] Because of the large number of atoms involved in polymeric methacrylic acids and large number of solvent molecules needed in the explicit solvation model, the implicit continuum solvation model were used in this work to determine the solvation free energies of the neutral acids and their conjugate bases. Currently there are several continuum solvation models available. In the so-called Solvation Model sets (SMx) developed by Cramer and Truhlar, the solvent free energies are determined by the bulk electrostatic interaction energies and surface tensions of the first shell solute molecules.[33,34] Another family of the continuum solvation models includes the original polarizable continuum model (PCM) developed by Tomasi et al.,[32] and later modified versions, the Conductor-like PCM (CPCM), isodensity PCM (IPCM), and self-consistent IPCM (SCIPCM).[35,36] In these models, the electrostatic interactions between the charges on the solute molecules and induced charges on the polarizable solvent molecules are solved self-consistently. These solvation models combined with structure optimization and atomic charge calculations using quantum mechanics yield solvation free energies in good agreement with available experimental data.[37] Furthermore, CPCM uses a simple scaled conductor boundary condition to improve its efficiency.[35,38] Liptack and Toth et al. have reported that the CPCM model is able to yield solvation free energies within 1 kcal/mol of the experiment values for the neutral acetic acid and acetate ion.[15,20] This model was used in this work to determine the solvation free energies of polymeric methacrylic acids and their base conjugates. It is worth mentioning that a hybrid approach, in which the first shell of the solvent is treated explicitly whereas the outer shells are treated with the implicit continuum model, also gives accurate solvation free energies.[16] We did not pursue this method in the current study due to the large number of solvent water molecules needed to solvate the methacrylic acid oligomers.

To the best of our knowledge, no theoretical calculations have been done to predict the pK_a values of nonbiological polymeric acids. In this work, we chose short methacrylic acid oligomers with DP from 1 to 3 (monomer, dimer, and trimer) to predict polymeric pK_a values and their dependence on the chain length. Our approach involves choosing proper thermodynamic cycles, selecting appropriate quantum mechanical methods and solvation models, and developing error reduction protocols.

2. Methodology

2.1. Thermodynamic Cycle. The value of pK_a determines the relative concentrations of an acid and its conjugate base at equilibrium. It directly relates to the free-energy change of the acid's proton dissociation process. Since it is difficult to directly calculate the free-energy change in solution (ΔG_{aq}), a thermodynamic cycle is generally employed to relate the acid, the conjugate base, and the proton in solution to the corresponding gas phase as shown in Scheme 1. The solution phase free-energy change ΔG_{aq} is directly related to the gas-phase free-energy change ΔG_{gas} of the same process and the solvation free energies of the reactants and products ΔG_{sol}

SCHEME 1: Thermodynamic Cycle for Determining the Free-Energy Change of Acid Proton Dissociation in Solution



$$\mathrm{p}K_a = \frac{1}{2.303RT}\left(\Delta G_{gas} - \Delta G_{sol}(\mathrm{HA}) + \Delta G_{sol}(\mathrm{A}^-) + \Delta G_{sol}(\mathrm{H}^+)\right) \quad (1)$$

It is possible to vary the chemical species involved in the cycle. For example, an acid can react with a water molecule to form a hydronium ion and a corresponding conjugate base or react with a hydroxyl ion to form a water molecule and the corresponding conjugate base. As long as the free energies are accurate, all the cycles should converge to the same pK_a value. The cycle presented above was used in the current work because it involves the least number of chemical species. This reduces the overall errors in the calculations of the gas-phase and solvation free energies. Recent studies on the performance of different thermodynamic cycles have shown that this simplest cycle gives the most accurate results.[17,20]

Alternatively a relative pK_a value can be obtained by taking the differences in the free energies between two thermodynamics cycles of similar acids (HA and HB)

$$\begin{aligned}\Delta \mathrm{p}K_a &= \mathrm{p}K_a(\mathrm{HB}) - \mathrm{p}K_a(\mathrm{HA}) \\ &= \frac{1}{2.303RT}\left(\Delta\Delta G_{gas}(\mathrm{HA}\rightarrow\mathrm{HB}) - \Delta\Delta G_{sol}(\mathrm{HA}\rightarrow\mathrm{HB}) + \Delta\Delta G_{sol}(\mathrm{A}^-\rightarrow\mathrm{B}^-)\right) \quad (2)\end{aligned}$$

This relative pK_a value by this two-cycle scheme should be less prone to the systematic errors in the free-energy calculations. Moreover, the solvation free energy for proton does not appear in the equation, thus eliminating the uncertainties in the experimental or calculated values.

2.2. Gas-Phase Free Energy. There are substantial improvements on the original complete basis set (CBS) model for gas-phase free-energy calculations.[27,39−41] In this work, we used the model developed by Petersson et al., the CBS-QB3, to calculate the gas-phase free energies for the acids and their conjugate bases. This model uses a CBS extrapolation of the electronic correlation energy based on the asymptotic convergence of pair orbital expansions to account for the truncation errors due to the finite size of the basis sets. Calculation of the CBS-QB3 free energies involve the following computational steps: (1) B3LYP/6-311G(2d,d,p) geometry optimization and zero-point energy (ZPE); (2) UMP2/6-311+G(3d2f,2df,2p) single-point energy and CBS extrapolation; (3) MP4(SQD)/6-31+G(d(f),p) energy; (4) CCSD(T)/6-31+G† energy. In addition, empirical parameters are used to scale ZPE and account for errors due to orbital overlap-interference and spin contamination.[39]

This CBS-QB3 model gives very good free-energy changes for over 50 molecular dissociation processes, with a mean absolute deviation of 0.87 kcal/mol.[39] However the very expensive high-order correlation calculations limit the application to small or moderately sized molecules with approximately less than 50 atoms. In this work, it was used for methacrylic acid and a few similar carboxylic acids in the gas phase. The less computationally demanding DFT based B3LYP functional was then used to reproduce these free energies. These two sets of free energies were compared to reveal the systematic errors involved in the DFT B3LYP calculations. The systematic errors in the free energies of gas-phase methacrylic acid short oligomers were then corrected.

2.3. Solvation Free Energy. Tomasi′s polarizable continuum model (PCM) is one of the most commonly used and reliable solvation models.[32,42,43] It treats the solvent as a continuum medium with an appropriate dielectric constant polarized by the charge and multipole distributions on the solute molecules. In our calculations, a modified PCM, the conductor-like PCM (CPCM) was used to calculate the solvation free energies. As mentioned earlier, this model uses a simpler boundary condition for the induced charges on the solute−solvent interface than the original PCM, thus more efficient in solving the reaction field energies self-consistently. It has been shown that it can predict solvation free energies for carboxylic acids and their conjugate bases to a high degree of accuracy.[15]

There are several factors that affect the PCM solvation free energies. The first one is the geometry of the solute. This geometry can be obtained by an optimization either in the gas phase or in solution. A previous investigation has shown that the solvation free energies for small neutral organic acids using the geometries optimized in the gas phase and in solution are comparable.[17] For their charged conjugate bases, solvation-free energies using the geometries optimized in solution agree better with the experiment data with an improvement of about ∼1 kcal/mol. In this work, the gas-phase-optimized structures were used for both the neutral acids and singly charged conjugate bases since only the relative solvation energies are important. The second factor is the choice of quantum methods and basis sets. Studies on CPCM model with various methods/basis sets have shown that higher-level theories and larger basis sets will not necessarily yield better solvation free energies.[15] In particular for carboxylic acids, the HF/6-31+G(d) was found to give the best results on the HF/6-31+G(d)-optimized gas-phase structures. The third factor is the atomic radii that determine the size and shape of the cavity where the solute is placed. Among available sets of atomic radii, we chose the UAHF set[37] over the UAKS set[44] in our calculations because it was developed specifically to be used with the HF method.

To use the one thermodynamic cycle as shown in Scheme 1, experimental solvation free energy for proton was used because accurate theoretical calculations are very difficult due to proton′s small radius and relative large charge. There are several experimental values available in literature for the proton solvation energy.[34,45−48] In our calcualations, we chose the value of −265.9 kcal/mol reported by Tissandies et al.[45] earlier and the most recent experiments by Kelly et al.[34] This solvation energy was determined based on aqueous solvation free energies of neutral acid species using the cluster pair approximation[45] and was considered to be the most accurate.[17,48] Because of the different standard states used, i.e., the experimental value is for a process from 1 atm in the gas phase to 1 mol/L in the solution at 298.15 K, whereas our standard states in both phases are 1 mol/L, the corrected experimental value should be −264.0 kcal/mol in our thermodynamic cycle as shown in Scheme 1.[49]

2.4. Calculation Procedure. The thermodynamic cycle shown in Scheme 1 was used to determine the pK_a values of acetic acid, propanoic acid, 2-methylpropanoic acid, pivalic acid (2,2-dimethylpropanoic acid), and monomer methacrylic acid (2-methyl-2-propenoic acid). All calculations were carried out using GAUSSIAN03 package.[50] Initially, the free energies of the acids and their conjugate bases in the gas phase were determined using the CBS-QB3 models. The starting structures were generated with various arrangements of the methyl groups and carboxyl groups. The gas-phase free energies of the various species were then recalculated using DFT with B3LYP/6-31+G(d) in an attempt to reproduce the CBS-QB3 free energies. The differences between the free energies from the two methods were determined and used to offset the systematic errors in DFT calculations for the pK_a calculations. The solvation free energies were calculated using the CPCM model with UAHF atomic radii based on HF/6-31+G(d)-optimized gas-phase structures.

The pK_a calculations for the dimer and trimer of methacrylic acid started with the gas-phase geometry optimizations with DFT B3LYP/6-31+G(d). The initial conformations of the dimer and trimer structures were generated by positioning the methyl and carboxyl groups of each repeating unit at various possible locations, except those with high steric repulsions between the neighboring methyl groups and carboxyl groups. Two optimized lowest-energy conformations (other conformations are at least 2 kcal/mol higher) were found for both dimer and trimer. These conformations were then used to generate the singly dissociated anions (two for the dimer and three for the trimer). The anions were also optimized with the same method in the gas phase. The gas-phase free energies were determined using vibrational analysis with B3LYP/6-31+G(d). The solvation free energies were determined using CPCM with the same procedure as described before. Our current work is not meant to give a comprehensive method for a broad selection of organic acids, but rather focuses on predicting pK_a values for short methacrylic acid oligomers.

3. Results and Discussion

3.1. Gas-Phase Free Energies. Before the pK_a values of the acid groups on methacrylic acid dimer and trimer were determined, several carboxylic acids with similar structures including acetic, propanoic, 2-methyl propanoic, pivalic, and methacrylic acids were used as a test set. The gas-phase free energies of the acids and their conjugate bases were calculated from their optimized gas-phase geometries. The starting geometries for optimization included all conformational possibilities, e.g., the hydrogen in the carboxyl group can take a trans or cis position to the acetyl oxygen, the methyl groups can be eclipsed or staggered to the carboxyl group. The CBS-QB3 free energies of these structures were then obtained from the lowest free energy conformations.

Table 1 lists the free-energy changes for acid dissociation in the gas phase given by the CBS-QB3 model. The experimental values are also listed for comparison.[51] For all the molecules tested, the CBS-QB3 model is able to predict free-energy changes that are within 1.9 kcal /mol of the experimental values. This error is comparable to the uncertainties found in the experiment (2.0−2.9 kcal/mol), indicating that the CBS-QB3 model performs well. However, this model cannot be used to determine the free energies of the methacrylic acid dimer and trimer because these molecules are prohibitively large for the

TABLE 1: Gas-Phase Free-Energy Changes for Proton Dissociation of Selected Carboxylic Acids Given by the CBS-QB3 Quantum Chemical Model and the Experiments (Uncertainties in Parentheses)

	acetic acid (kcal/mol)	propanoic acid (kcal/mol)	2-methyl propanoic acid (kcal/mol)	pivalic acid (kcal/mol)	methacrylic acid (kcal/mol)
experiment	341.1 (2.0)	340.3 (2.0)	338.9 (2.0)	337.7 (2.0)	337.0 (2.9)
CBS-QB3	339.48	340.13	336.99	336.49	335.88
SCF[a]	355.37	355.44	354.24	353.91	351.92
ZPE[b]	346.68	346.66	345.44	345.02	343.20
THC[c]	346.49	347.73	345.15	344.98	343.15
MP2[d]	338.14	338.79	335.56	335.01	334.36
CBSE[e]	336.34	337.00	333.76	333.19	332.56
MP4[f]	339.74	340.44	337.35	336.86	336.02
CCSD[g]	338.68	339.31	336.13	335.59	335.05
SPIN[h]	339.34	339.97	336.81	336.29	335.72
EMP[i]	339.48	340.13	336.99	336.49	335.88

[a] SCF HF single-point energy. [b] Zero-point energy. [c] Thermal corrections. [d] MP2 energy. [e] Complete basis set extrapolation. [f] MP4 energy. [g] CCSD energy. [h] Spin contamination correction. [i] Empirical corrections. The CBS-QB3 calculation details are given in Reference 39.

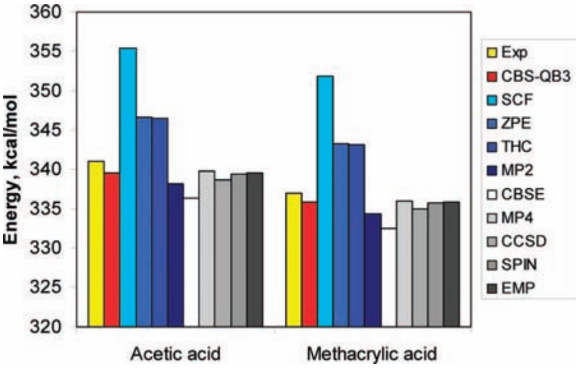


Figure 1. Experiment and CBS-QB3 gas-phase free-energy changes for the proton dissociation processes of acetic acid and methacrylic acid. The free-energy changes by progressively adding corrections from each of the calculations in the CBS-QB3 model are shown in different shades of blue and gray.

higher order calculations such as MP4 and CCSD used in the model. A less computationally demanding method needs to be used other than CBS-QB3. Also listed in Table 1 are the CBS-QB3 free-energy changes by progressively adding the corrections from each of the calculations in the model. The relatively efficient calculations of the DFT with B3LYP functional and the MP2 single-point energy are sufficient to bring the free energies to about 3.3 kcal/mol within the experiment values for all the acids tested. On the other hand, the demanding MP4 and CCSD calculations, and CBS extrapolation are only able to improve the free energies by an additional 0.7 kcal/mol on average. The rest of the corrections come from correction for spin contamination and an additional empirical term.[40] More importantly, the combined contributions by CBS extrapolation, MP4, CCSD, and spin and empirical corrections remain more or less constant at −1.42 ± 0.08 kcal/mol for the carboxylic acids tested.

Figure 1 plots the free energies of acetic and methacrylic acids with various corrections. Our results suggest that the higher order calculations may be avoided; instead, the free-energy changes can be obtained by simply applying a correction to the efficient SCF, ZPF, THC and MP2 calculations without sacrificing much of the accuracy. This approximation is particularly desired for the large structures of dimer and trimer of the methacrylic acid. Further, DFT with the modified Becke hybrid functional B3LYP was used to determine the gas-phase free energies instead of the MP2 method.[28,29] DFT with B3LYP functional was chosen because it has superior ratio of accuracy to cost and is less prone to spin contamination than Møller−Plesset perturbation theory. B3LYP functional has been used in several versions of the accurate Gaussian-*n* and CBS models (as seen in CBS-QB3). For the basis set we chose the moderate 6-31+G(d). Table 2 shows the gas-phase free-energy changes for the dissociation processes of the small carboxylic acids calculated by CBS-QB3, MP2 in CBS-QB3, and B3LYP/6-31+G(d). The B3LYP method with our basis set underestimates the free energy by about −4.88 ± 0.73 kcal/mol on average. This amount of correction can be added to the B3LYP/6-31+G(d) free energies to reproduce the CBS-QB3 free energies at a much lower computational cost than the CBS-QB3 model.

3.2. Solvation Free Energy. The other quantities involved in the thermodynamic cycle are the solvation free energies for the acids, their corresponding conjugate bases, and the proton. The CPCM solvation model was used in our calculations for the acids and their conjugate bases. Experimental proton solvation free energy of −264.0 kcal/mol was used.[45] Liptack et al. has shown that CPCM gives more accurate solvation free energies for carboxylic acids than other commonly used continuum models.[15] The same authors also investigated the effects of the level of theory and the size of basis set on solvation free energy. They found that high level theories and large basis sets will not necessarily yield more accurate results. The simple HF/6-31+G(d) is sufficient to give solvation free energies within 2 kcal/mol of the experiment value. This finding is confirmed in another solvation free-energy study.[16] In the current work, HF/6-31+G(d) was first used to optimize the gas-phase geometries. The optimized structures were then used in the solution phase to determine the solvation free energies.

Table 3 lists the solvation free energies given by the CPCM model. For the acetic acid and its conjugate anion, experiment solvation free energies are −6.70 and −77.3 kcal/mol, respectively.[52] Our calculated results of −7.8 and −75.5 kcal/mol agree well with the experimental data with errors less than 2 kcal/mol. It is difficult to assess the systematic errors in the calculated CPCM solvation free energies for all the acids tested due to the lack of experimental data for comparison. The solvation free energies calculated by PCM models depend critically on the size and shape of the solute cavity (which is built by interlocking spheres centered at the atoms) and the dielectric constant of the solvent. By use of the same UAHF set of atomic radii and a dielectric constant of 78.39 for water for the all calculations, systematic errors are minimized. Kelly et al. reported an experimental solvation free energy of −76.2 kcal/mol for the propanoic acid anion.[34] The free energy from the current CPCM calculation is 1.9 kcal/mol higher, comparable to error for the acetic acid anion.

TABLE 2: Comparison of Gas-Phase Free-Energy Changes for Proton Dissociation of Small Carboxylic Acids given by CBS-QB3, MP2, and DFT B3LYP Calculations

	acetic acid (kcal /mol)	propanoic acid (kcal/mol)	2-methyl propanoic acid (kcal/mol)	pivalic acid (kcal/mol)	methacrylic acid (kcal/mol)
ΔG_{gas}(CBS-QB3)	339.48	340.13	336.99	336.49	335.88
ΔG_{gas}(MP2)	338.14	338.79	335.56	335.01	334.36
$\Delta\Delta G_{gas}$ [a]	−1.34	−1.34	−1.43	−1.48	−1.52
ΔG_g(B3LYP)	334.47	334.03	332.66	332.05	331.34
$\Delta\Delta G_{gas}$ [b]	−5.01	−6.10	−4.32	−4.44	−4.54

[a] Mean deviation with respect to CBS-QB3 is −1.42 ± 0.08 kcal/mol. [b] Mean deviation with respect to CBS-QB3 is −4.88 ± 0.73 kcal/mol.

TABLE 3: Solvation Free Energies of Small Carboxylic Acids (Neutral) and Their Conjugate Bases (Anion) Given by the CPCM Solvation Model with HF/6-31+G(d)

solvation free energy (kcal/mol)		acetic acid	propanoic acid	2-methyl propanoic acid	pivalic acid	methacrylic acid
neutral acid	CPCM with HF/6-31+G(d)	−7.83	−7.91	−7.31	−6.52	−6.01
	experiment	−6.70[a]				
anion	CPCM with HF/6-31+G(d)	−75.53	−74.26	−73.02	−70.12	−69.89
	experiment	−77.3[a]	−76.2[b]			

[a] Reference 48. [b] Reference 34.

TABLE 4: Experimental and Calculated pK_a Values for Small Carboxylic Acids with CBS-QB3 and B3LYP Gas-Phase Free Energies[a]

	experimental	CBS-QB3	CBS-QB3/empirical correction	B3LYP	B3LYP/empirical correction	B3LYP/two-cycle
acetic acid	4.75	5.70	4.66	5.61	4.61	4.75
propanoic acid	4.87	7.17	5.24	6.27	4.89	5.41
2-methyl propanoic acid	4.83	5.34	4.52	5.74	4.67	4.88
pivalic acid	5.03	6.52	4.99	6.84	5.13	5.98
methacrylic acid	4.65	5.86	4.72	6.11	4.83	5.26
mean absolute deviation		1.29	0.18	1.29	0.12	0.54

[a] The solvation free energies are given by the CPCM model in all pK_a calculations. The pK_a values corrected using the linear empirical correlation and eq 2 (in Section 2) are also shown.

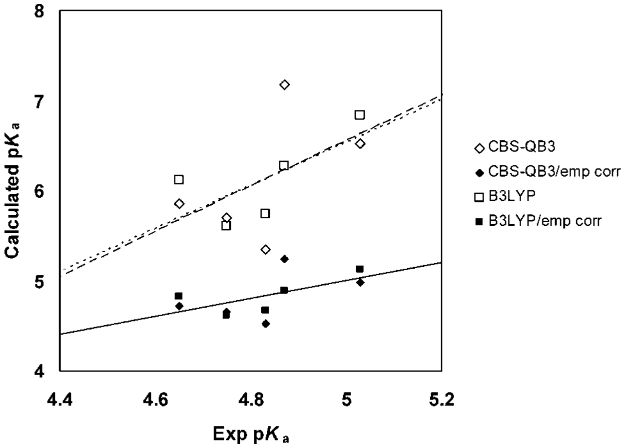


Figure 2. Comparison of experimental and calculated pK_a values for the small carboxylic acids. The dashed line shows the linear fit to the CBS-QB3 derived pK_a; the dotted line to the B3LYP derived pK_a; the solid line represents the perfect agreement between the calculated and experimental pK_a.

3.3. pK_a Values for Carboxylic Acids. The calculated pK_a values for the small carboxylic acids with CBS-QB3 gas-phase free energies and the CPCM solvation free energies are listed in Table 4 and plotted in Figure 2. The mean absolute deviation from experimental data is 1.29 pK_a units for these acids. It was shown previously that the B3LYP/6-31+G(d) method with a correction of 4.88 kcal/mol can reproduce the CBS-QB3 free energies. Table 4 also lists the calculated pK_a values using this approach with the same CPCM solvation free energies. The mean absolute deviation is again 1.29 pK_a units, indicating that the much faster DFT B3LYP method can replace the very expensive CBS-QB3 model without lowering accuracy in the predicted pK_a values.

The pK_a values of the selected acids were also determined using the two-cycle scheme with gas-phase free energies determined from DFT B3LYP to reduce systematic errors in the calculations. The thermodynamic cycle of acetic acid was used as the standard, on which all the other acids were based. The experimental acetic acid pK_a value of 4.75 was used as pK_a (HA) in eq 2. The calculated pK_a (HB) values are listed in Table 4. An absolute mean pK_a deviation of 0.54 was obtained.

To further improve the predicted pK_a values, an empirical correction suggested by Klicic et al. was used.[13] These authors studied a wide range of organic compounds and used a simple linear relationship, $\mathrm{p}K_a = A(\mathrm{p}K_a^{\mathrm{cal}}) + B$, to correct the calculated pK_a values, where A and B are fitted constants. For a total of 40 carboxylic acids with pK_a values ranging from −0.4 to 5.0, an A of 0.4035 and B of 0.155 were obtained. The predicted pK_a values have a mean absolute deviation of 0.5 pK_a unit. The same formula was used to improve calculated pK_a values in this work. The results are listed in Table 4. For the pK_a values calculated using the CBS-QB3 gas-phase model, parameter A is found to be 0.397 and B to be 2.40. This reduces the mean absolute pK_a deviation to 0.18 pK_a units. For the pK_a values calculated using DFT B3LYP method, the parameter A is found to be 0.418 and B to be 2.27. The mean absolute pK_a deviation is only 0.12 pK_a units. The differences between our and Klicic′s A and B values are due to different quantum chemical theories, basis sets, and solvation models and atomic radii used in the calculations.

The calculated free energies from CBS-QB3 and DFT B3LYP with CPCM solvation model tend to overestimate the pK_a values

TABLE 5: Gas-Phase and Solvation Free Energies for the Two Lowest-Energy Conformations of Methacrylic Acid Dimer and Trimer and Their Corresponding Calculated pK_a Values for Each Carboxyl Group

	gas-phase free energy (B3LYP), kcal/mol		solvation free energy (CPCM), kcal/mol		pK_a		
	neutral acid	anion	neutral acid	anion	B3LYP	B3LYP/empirical correction	B3LYP/two-cycle
methacrylic acid dimer (D1)	−385329.93	−384996.67	−13.27	−69.85	8.27	6.08	6.98
	−385329.93	−384995.98	−13.27	−73.20	6.32	5.27	5.03
methacrylic acid dimer (D2)	−385329.07	−384995.01	−12.86	−72.46	6.64	5.41	5.36
	−385329.07	−384996.32	−12.86	−71.11	6.67	5.42	5.39
methacrylic acid trimer (T1)	−577609.69	−577278.25	−15.27	−72.07	6.77	5.46	5.49
	−577609.69	−577281.50	−15.27	−68.24	7.20	5.64	5.91
	−577609.69	−577275.04	−15.27	−75.59	6.55	5.37	5.26
methacrylic acid trimer (T2)	−577609.61	−577276.57	−16.59	−74.07	7.45	5.74	6.16
	−577609.61	−577280.08	−16.59	−72.01	6.38	5.30	5.10
	−577609.61	−577275.44	−16.59	−75.65	7.12	5.60	5.83

by 0.5−2.3 units as shown from Table 4. The CBS-QB3 model, based on the complete basis set extrapolation, can predict the free-energy changes in the gas phase within 1.9 kcal/mol of the experimental values, which is comparable to the experimental accuracy. This amount of error corresponds to 1.4 pK_a units. The average error involved in the CPCM solvation free-energy change for proton dissociation of the acids is about 3 kcal/mol or another 2.2 pK_a units. Therefore, the mean absolute error of only 1.29 pK_a units given by the calculations is likely due to the partial cancelation of errors. This is easily seen by examining the deviations of the calculated free energies. As shown in Table 1, all CBS-QB3 free-energy changes are systematically lower than the experimental values. This negative deviation in ΔG_{gas} contributes negatively to the pK_a value (see eq 1). On the other hand, the CPCM solvation free energy is 1.1 kcal/mol lower than the experimental value for the acetic acid and 1.8 kcal/mol higher for the acetate ion. These deviations all contribute positively to the pK_a values and partially cancel the negative contribution from the gas-phase calculations.

In the solvation free-energy calculations, the errors come mainly from the constructed cavity that may not truly represent the solute in solution and the neglect of the explicit solvent molecules that form hydrogen bonds with the solute molecules. In our CPCM calculations, the cavity is constructed using interlocking spheres centered at the atoms with the UAHF set of atomic radii. This cavity may be different from what the solute actually occupies in the solvent. As a result, the charge distribution on the solute molecule is different and also the response of the polarizable solvents to the charge distribution. All these affect the electrostatic interaction energy which is one of the major contributions to the solvation free energy. The neglect of explicit hydrogen bonding interaction between solute and solvent molecules will also affect the solvation free energy.

The reason that an empirical correction with linear relationship $\mathrm{p}K_a = A(\mathrm{p}K_a^{cal}) + B$ works can be hinted from eq 2, employing the two-thermodynamics cycle scheme. Equation 2 indicates that the pK_a values can be systematically shifted to fit best with the experimental data and cancel out some of the systematic errors inherent in the calculations. This accounts for the *B* parameter. Earlier work by Klamt and co-workers found that the slope *A* is in the range of 0.6.[12] Chipman and Klicic et al. found the slope *A* to be 0.47 for acids.[7,13] Our values are close to 0.40, similar to their results. It is perplexing that the slope *A* is not close to 1 as found by Klamt and co-workers.[12] Further theoretical study is required to reconcile this difference.

3.4. pK_a Values for Methacrylic Acid Dimer and Trimer. In the previous sections, it has been shown that in the protocol developed for small carboxylic acids B3LYP/6-31+G(d) is able to yield good gas-phase free energy if a constant offset is applied. The CPCM solvation model, though it neglects the explicit solvent molecules and hydrogen bonding interaction between the solute and solvent molecules, is still able to give relatively accurate solvation free energies. By use of the linear empirical correction, the predicted pK_a values are in excellent agreement with the experimental values. The same protocol was applied to the dimer and trimer of methacrylic acid, in hoping that the protocol's ability to well account for the errors in the small carboxylic acids is retained in the short oligomers of methacrylic acid.

The conformational structures of the dimer and trimer of methacrylic acid will affect the pK_a values. As mentioned in section 2, the conformations with the lowest gas-phase free energies were used in the CPCM model to calculate the solvation free energies for the small monomeric carboxylic acids. For large organic molecules or polymers, their conformations in solution will be different from their gas-phase structures due to hydrophobic and/or hydrogen bonding interactions. In this study, the conformational structures of the dimer and trimer of methacrylic acid are assumed to have extended chains. Methacrylic acid is rather hydrophilic. Open chain structures are expected for short oligomers of methacrylic acid in aqueous solution.

The gas-phase free energies for the two lowest-energy dimer and trimer conformations of methacrylic acid are listed in Table 5. The free energies given by other conformations are at least 2 kcal/mol higher. At 298 K, this energy difference means that only 3.5% of the total population will take the high-energy conformations so they can be safely removed from further consideration. The two lowest energy conformations for the dimer and trimer of methacrylic acid are shown in Figure 3. Since only the first proton dissociation constants are of interest here, the singly charged anions (two anions from the dimer and three from the trimer) were generated by removing the proton from each of the carboxyl groups and their gas-phase free energies are listed in Table 5. The CPCM at HF/6-31+G(d) solvation free energies are also listed in Table 5.

The predicted pK_a values in Table 5 using eq 1 show an overall increase over the calculated pK_a value of 6.11 for the monomer (see Figure 4) by using the B3LYP gas-phase free energies. For instance, the lowest energy conformation D1 of the dimer gives 8.27 and 6.32 for the two carboxyl groups. This discrepancy between these two pK_a values reflects different environments for the two carboxyls in this conformation. However, as seen in the second lowest energy conformation D2, the pK_a values are very similar with 6.64 and 6.67, respectively. After the linear empirical correction using the correlation developed for the tested carboxylic acids, the pK_a

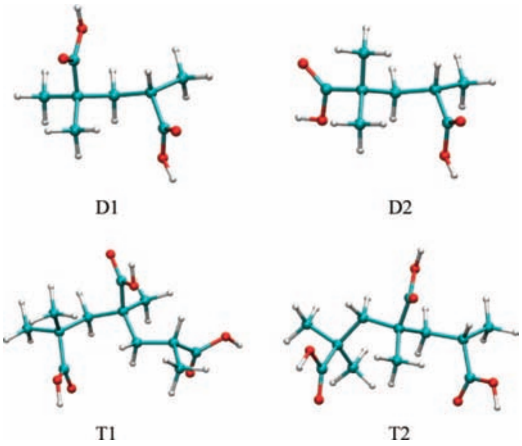


Figure 3. Lowest free-energy conformations for the methacrylic acid dimer and trimer in the gas phase. D1, dimer with the lowest free energy; D2, dimer the second lowest; T1, trimer the lowest; T2, trimer the second lowest. Their free energies are given in Table 5.

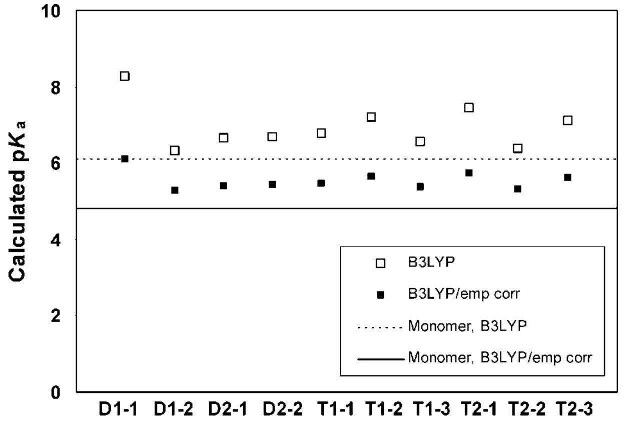


Figure 4. Calculated pK_a values for the methacrylic acid dimer and trimer and a comparison with the corresponding calculated monomer pK_a values. The abbreviation D1−1 on the X-axis is for the dimer, conformation 1 (with the lowest free energy), and the first COOH group.

values become 6.08, 5.27, 5.41, and 5.42, respectively, with 0.4−1.25 pK_a units higher than the corrected theoretical monomer value of 4.83 (Figure 4). For the methacrylic trimer, the pK_a values for the lowest energy conformation T1 are 6.77, 7.20, and 6.55 for the three carboxyl groups, respectively. After the empirical correction, the corresponding pK_a values become 5.46, 5.64, and 5.37. The pK_a values for the second lowest energy conformation T2 are 7.45, 6.38, and 7.12 for the three acid sites respectively without empirical correction. They are 5.74, 5.30, and 5.60 with correction. Again, these values are at least 0.5 pK_a units higher than the corrected theoretical monomer pK_a value. The mean deviations of dimer and trimer′s pK_a values from the monomer′s pK_a are 0.87 and 0.80 before the empirical correction and 0.71 and 0.69 after the correction, respectively. Alternatively, the ΔpK_a values with respect to the monomer methacrylic acid can be calculated from eq 2 (section 2) using two cycle scheme, where the calculated pK_a value for the monomer (4.83) is used as the pK_a(HA) in eq 2. The pK_a values using this two-cycle scheme are listed in Table 5. For all carboxyl groups in the dimer and trimer, the pK_a values again show an increase over the monomer value.

This increase in pK_a value is found for every carboxyl group in the methacrylic acid dimer and trimer, indicating that the first proton dissociation process of these oligomers becomes less favorable upon polymerization. The differences in pK_a values between the carboxyl groups in dimer and trimer are small and negligible.

Experimental evidence suggests that the pK_a values of the polymeric methacrylic acids could be significantly larger than the corresponding monomer value.[53] Our calculated pK_a values for the methacrylic acid dimer and trimer indicate the same trend. The theoretical pK_a values of the dimer and trimer of methacrylic acids are found to be higher than the theoretical monomer methacrylic acid pK_a value by about 0.7 pK_a units. The reason is that deprotonated methacrylic acid monomer forms a large conjugated delocalized π bond thereby stabilizing the structure, whereas there is no conjugated π bond formation after proton dissociation from the carboxyl groups in the dimer and trimer.

4. Conclusions

We have presented a protocol to predict the pK_a values for small carboxylic acids via combined quantum mechanical calculations and continuum solvation theory using thermodynamics cycles. The free energies obtained from the complete basis set calculations with CBS-QB3 agree very well with the experimental values. However, it is not suitable for our target molecules, methacrylic acid dimer, and trimer due to the large number of atoms involved. Instead, we used the DFT with B3LYP functional and the 6-31+G(d) basis set to reproduce the gas-phase free energies after a correction. The solvation free energies of the acids and their conjugate bases were given by CPCM using gas-phase geometries optimized using HF/6-31+G(d). The pK_a values calculated are slightly higher than the corresponding experimental data with a mean absolute deviation of 1.29 pK_a units. This deviation was greatly reduced to 0.12 pK_a units after applying a linear empirical correction.

This protocol was used to determine the pK_a values of the first proton dissociation constants for all the carboxyl groups in the methacrylic acid dimer and trimer. For both dimer and trimer, the predicted pK_a values are on an average 0.7−0.8 pK_a units higher than the theoretical pK_a value of the monomer. This result indicates that carboxyl groups become less acidic after polymerization due mainly to the loss of π conjugation.

Acknowledgment. This work was supported by the National Science Foundation under CBET-0651231. Helpful discussions with Dr. Ranil Wickramasinghe from Colorado State University and Dr. Scott Husson from Clemson University are gratefully acknowledged.

References and Notes

(1) Bell, R. P. *The Proton in Chemistry*; Chapman and Hall: London, 1973.

(2) Forsyth, W. R.; Antosiewicz, J. M.; Robertson, A. D. *Prot.-Struct., Funct. Genet.* **2002**, *48*, 388.

(3) Mbaraka, I. K.; Shanks, B. H. *J. Catal.* **2006**, *244*, 78.

(4) Baker, R. W. *Membrane Technology Applications*; Wiley: New York, 2004.

(5) Wang, J. S.; Matyjaszewski, K. *J. Am. Chem. Soc.* **1995**, *117*, 5614.

(6) Chen, I. J.; MacKerell, A. D. J. *Theor. Chem. Acc.* **2000**, *103*, 483.

(7) Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413.

(8) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **1999**, *103*, 11194.

(9) Gao, D.; Svoronos, P.; Wong, P. K.; Maddalena, D.; Hwang, J.; Walker, H. *J. Phys. Chem. A* **2005**, *109*, 10776.

(10) Hudaky, P.; Perczel, A. *J. Phys. Chem. A* **2004**, *108*, 6195.

(11) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 2493.

(12) Klamt, A.; Eckert, F.; Diedenhofen, M.; Beck, M. E. *J. Phys. Chem. A* **2003**, *107*, 9380.

(13) Klicic, J. J.; Friesner, R. A.; Liu, S. Y.; Guida, W. C. *J. Phys. Chem. A* **2002**, *106*, 1327.

(14) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610.

(15) Liptak, M. D.; Shields, G. C. *Int. J. Quantum Chem.* **2001**, *85*, 727.

(16) Pliego, J. R. J.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *106*, 7434.

(17) Sadlej-Sosnowska, N. *Theor. Chem. Acc.* **2007**, *118*, 281.

(18) Shuurmann, G.; Cossi, M.; Barone, V.; Tomasi, J. *J. Phys. Chem. A* **1998**, *102*, 6706.

(19) Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A. *J. Phys. Chem. A* **1997**, *101*, 10075.

(20) Toth, A. M.; Liptak, M. D.; Phillips, D. L.; Shields, G. C. *J. Chem. Phys.* **2001**, *114*, 4595.

(21) Tunon, I.; Silla, E.; Tomasi, J. *J. Phys. Chem.* **1992**, *96*, 9043.

(22) Yang, P.; Murthy, P. P.; Brown, R. E. *J. Am. Chem. Soc.* **2005**, *127*, 15848.

(23) Davies, J. E.; Doltsinis, N. L.; Kirby, A. J.; Roussev, C. D.; Sprik, M. *J. Am. Chem. Soc.* **2002**, *124*, 6594.

(24) Tang, C. L.; Alexov, E. G.; Pyle, A. M.; Honig, B. *J. Mol. Biol.* **2007**, *366*, 1475.

(25) Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press: Oxford, 1987.

(26) Curtiss, L. A.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.

(27) Montgomery, J. A. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1994**, *101*, 5900.

(28) Becke, A. D. *J. Chem. Phys.* **1992**, *96*, 2155.

(29) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(30) Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc.* **1987**, *109*, 6857.

(31) Li, J.; Fisher, L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35*, 4694.

(32) Miertus, S.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1981**, *55*, 117.

(33) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.

(34) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 16066.

(35) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *24*, 669.

(36) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098.

(37) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210.

(38) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799.

(39) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822.

(40) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. *J. Chem. Phys.* **1996**, *104*, 2598.

(41) Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991**, *94*, 6081.

(42) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253.

(43) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43.

(44) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

(45) Tissandies, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V. *J. Phys. Chem. A* **1998**, *102*, 7787.

(46) Zhan, C. G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534.

(47) Palascak, M. W.; Shields, G. C. *J. Phys. Chem. A* **2004**, *108*, 3692.

(48) Camaioni, D. M.; Schwerdtfeger, C. A. *J. Phys. Chem. A* **2005**, *109*, 10795.

(49) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row, Publishers, Inc: New York, 1976.

(50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J.C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al.Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Wallingford CT, 2004.

(51) Burgess, D. R. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 20899, http://webbook.nist.gov.

(52) Pliego, J. R. J.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622.

(53) Tsujii, Y.; Hirose, Y.; Ejaz, M.; Fukuda, T.; Ishidoya, M. *Polymer Preprint (Am. Chem. Soc., Div. Polym. Chem.)* **2002**, *43*, 317.

JP807315P