JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Explicit Solvation Shell Model and Continuum Solvation Models for Solvation Energy and pK_a Determination of Amino Acids

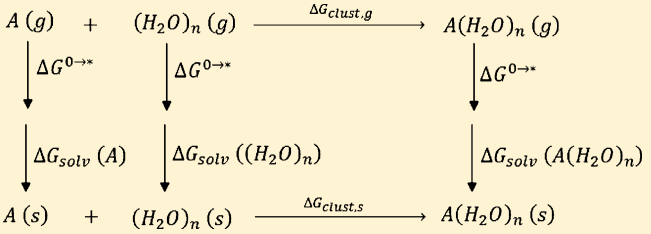
Mayuri Gupta,[†] Eirik F. da Silva,[‡] and Hallvard F. Svendsen*,[†]

[†]Department of Chemical Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway

[‡]Department of Process Technology, SINTEF Materials and Chemistry, N-7465 Trondheim, Norway

Supporting Information

ABSTRACT: The study of the Explicit Solvation Shell Model (ESS) presented recently [da Silva, E. F.; Svendsen, H. F.; Merz, K. M. *J. Phys. Chem. A* **2009**, *113*, 6404.] for calculation of solvation free energy of ions is extended for the study for amino acids. Solvation free energies and pK_a of a data set of 10 amino acids is calculated using ESS. The data set of amino acids is selected based on their potential to be regarded as solvents for postcombustion CO_2 capture processes. Calculated results are compared against experimental pK_a and pK_a calculated from PCM, SM8T, and DivCon continuum solvation models. Error estimates of pK_a from different models vs experimental pK_a data are also given to evaluate the results calculated by different solvation models. This study also involves a comprehensive study of gas phase basicity, proton affinity,ΔG^0_{acid}, ΔH^0_{acid}, protonation entropy with density functional methods (B3LYP/6-311++G(d,p)) and composite methods (G3MP2B3, G3MP2, CBS-QB3, G4MP2) and their comparison with experimental results for amino acids.



1. INTRODUCTION

The accurate prediction of solvation free energies of ionic species and their pK_a by employing appropriate thermodynamic cycles is of crucial importance in many areas of chemistry and biochemistry.[1,2] In spite of the immense importance of solvation free energies, accurate experimental results are available for only a few hundred small molecules of less chemical complexity and diversity. Also, current experimental techniques, dating back to Wolfenden and others,[3−6] involve accurate measurements of vapor pressure. In addition, very low vacuum concentrations and solubility issues of molecules with large negative solvation energies make them difficult to measure in these experiments. Because of all these experimental limitations, easily measurable and available solvation free energies are only for the range of +4 kcal/mol (insoluble molecules) to −11 kcals/mol (strongly solvated molecules).

Until now, it has not been possible to measure experimentally free energies of amino acids[7] and precise computational prediction of free energies of solvation is only possible for small molecules, as explained above. The scarcity of experimental data and limitations of experimental techniques make the computational prediction of solvation free energy very challenging. Various computational approaches have been suggested to calculate free energy of solvation of amino acids such as the group additivity methods[8,9] of the amino acid side chains (side chain analogs, e.g., methanol for Serine, etc.).[6] Konig et al.[10] observed an error of 6.7 kcal/mol from their study of blocked amino acids, and they suggested that it is difficult to meet Dill's criterion,[11] that is, an error of 1 kcal/mol or less for a full protein with additivity-based methods. Addivitity methods of predicting solvation free energy of amino acids from side chain analogs will be having average errors of typically 400 kcal/mol for a hypothetical protein containing 100 amino acids (not considering the errors of the secondary structure or tertiary structure contacts, which cannot be accounted for in a group-additive model). Such errors are clearly unacceptable. Hence, for accurate calculation of solvation free energies, the use of free energy simulations with explicit or generalized Born-based implicit solvent models[12,13] is a more promising approach available at present for amino acids and structured biomolecules as studied in the present work.

The objective of this work is to extend our study employing the explicit solvation shell model and present a prediction of reasonably good solvation energies for amino acid ions and to compare the model performance with the existing continuum solvation models. In the literature, various examples such as solvent reaction fields applied in a self-consistent manner to Schrödinger's equation polarizable continuum model (PCM),[14] Poisson−Boltzmann (PB) theory,[15] group additivity methods,[16−19] and chemical typing of surface area[20−22] as well as various combinations thereof[23,24] are available for calculating free energies of solvation of amino acids.

In the present work, solvation free energy of amino acids is studied because of two main motivations. First, the experimental free energy of amino acids is not available and as these are highly charged species (anionic, cationic, zwitterionic, and dianionic), this provided us with a possibility of comparing the performance of various continuum solvation models with the explicit solvation shell model. Second, the set of amino acids studied in this work is considered to be potential solvents for postcombustion CO_2

Received: June 2, 2013
Published: September 27, 2013

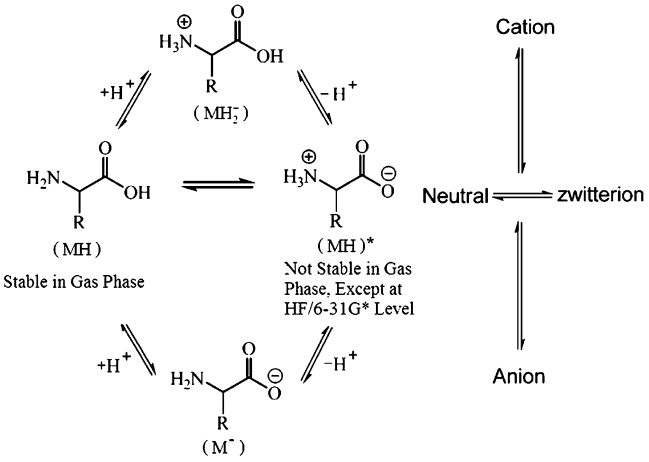
   dx.doi.org/10.1021/ct400459y | *J. Chem. Theory Comput.* 2013, 9, 5021−5037

capture processes and our results would provide interesting insights for future solvent development for PCC.

Solvation free energies of small neutral molecules can be calculated by using several quantum-continuum models within experimental error bars of ~1 kcal/mol.[14,25,26] However, precise prediction of the solvation free energies for ions with continuum dielectric solvent models that are not parametrized for the same is a challenging task when we are dealing with ionic solutes having concentrated charge densities with strong local solute−solvent interactions.[27−29] Since continuum methods cannot accurately describe short-range intermolecular interactions such as hydrogen bonding, it has become common practice to add explicit solvent molecules to the model ionic systems,[30−49] for example, in cluster-continuum studies of ion hydration free energies[20,34,38,39,43,47−53] and for pK_a calculations.[35,40,54−56] For example, Kelly et al. showed that adding one explicit solvent molecule significantly improves the accuracy of the calculated aqueous solvation free energies[21,34] and the acid dissociation constants[35] when using the SM6 continuum solvation model. Pliego and Riveros reported that a cluster-continuum solvent model with two to three explicit water molecules gives calculated pK_a's of 17 organic molecules that are in better agreement with the measurements than those calculated by using pure continuum solvent methods.[39,40] Some calculations have also shown that solvation calculations for bare metal ions should include at least a complete first coordination sphere of solvent (water) molecules.[30−32,36,37,41,42,44−46]

Here, we present a study of free energy of solvation of the 10 amino acids calculated using the explicit solvation shell model

Scheme 1. Gas Phase Protonation and Deprotonation Reaction Equilibria for Amino Acids



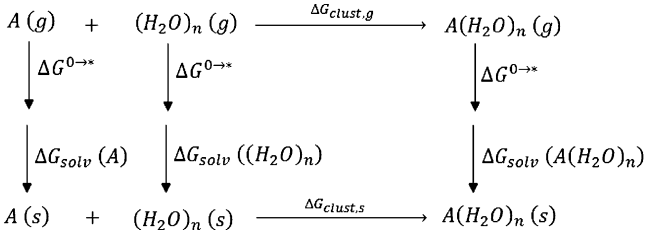


Figure 1. Thermodynamical cycle for computing solvation free energies with explicit solvation shell model.

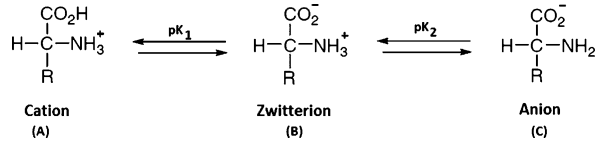


Figure 2. Equilibrium of different structures in the calculation of pK_{a1} and pK_{a2} of amino acids.

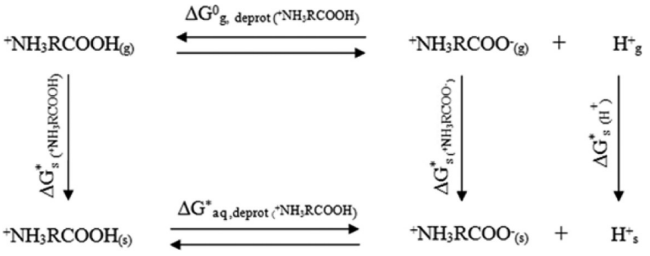


Figure 4. Thermodynamical cycle employed for calculation of pK_{a1} values.

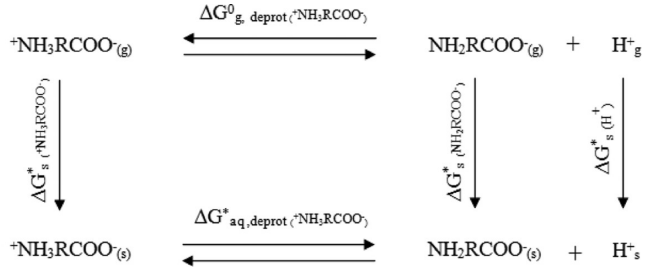


Figure 5. Thermodynamical cycle employed for calculation of pK_{a2} values.

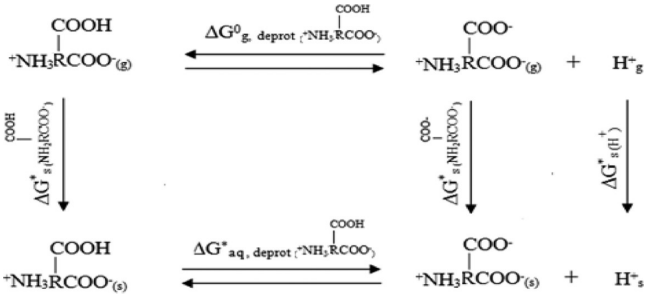


Figure 6. Thermodynamical cycle employed for calculation of pK_{a3} values.

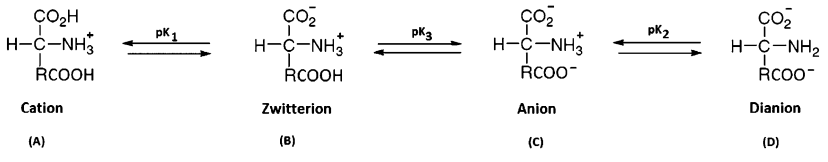


Figure 3. Equilibrium of different structures in the calculation of pK_{a1}, pK_{a2}, and pK_{a3} of amino acids.

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

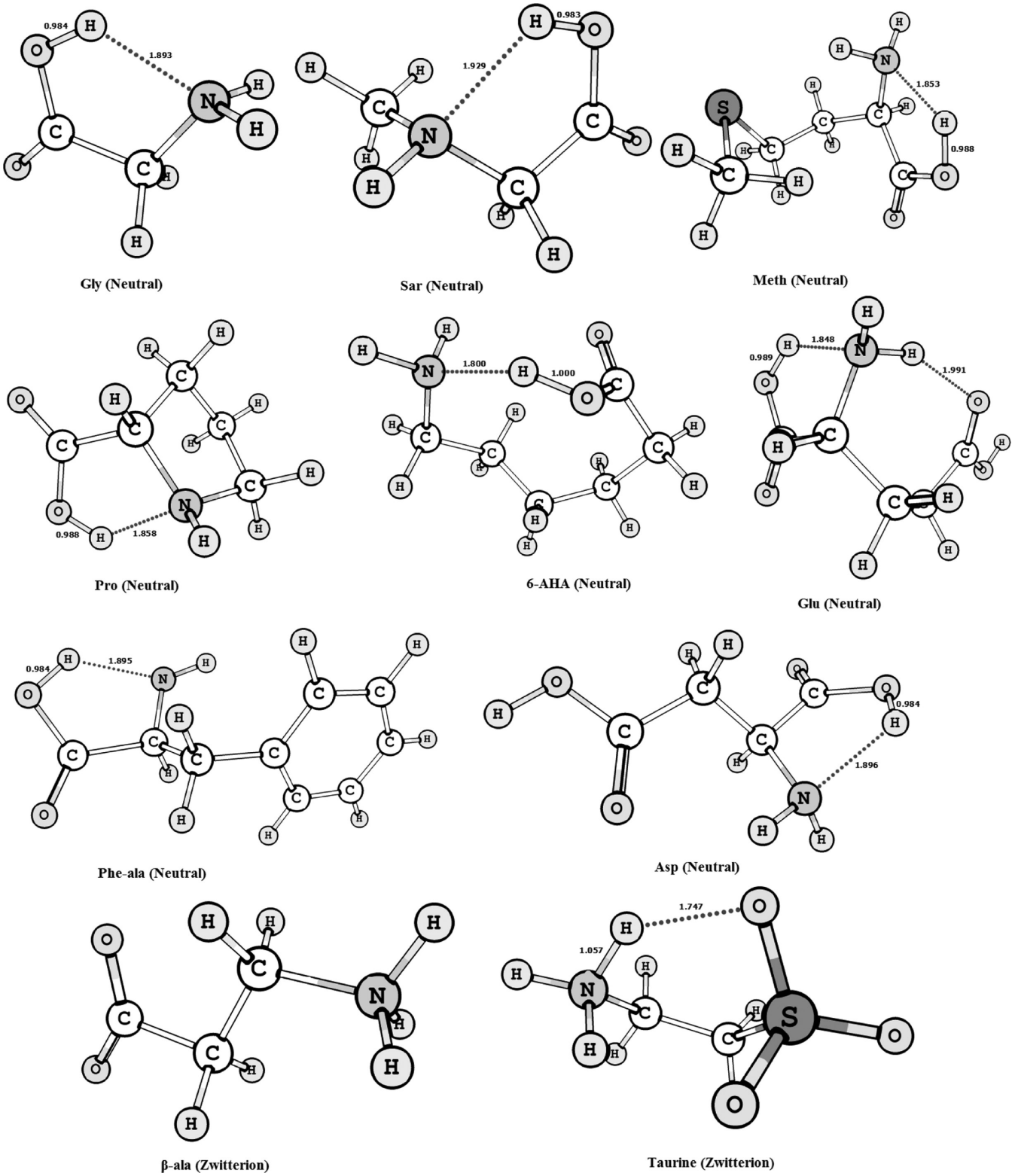


Figure 7. Optimized gas phase (neutral ⇆ zwitterion) form of amino acids studied in present work at CBS-QB3 level of theory. (The distance between amino group N−H and carboxyl group O−H is given in Angstrom.)

(ESS)[51] and three continuum solvation models (PCM, SM8T, DivCon). More specifically, we report (a) free energy of solvation of amino acids at pH 2−3, in their cationic form (b) free energies of solvation of amino acids at pH 7, in their zwitterionic forms, (c) free energies of solvation of amino acid side chains (d) Amino group pK_a, Carboxyl group pK_a and amino acid side chain pK_a at 298 K, where pK_a results are compared with available experimental data, and, finally, (e) a comparison of various pK_a and solvation free energy of ions calculated from explicit solvation shell and continuum solvation shell models (polarized continuum solvation shell model, SM8T, Divcon).

In this work, gas phase thermochemical properties (i.e., gas phase basicity, proton affinity, acidity, deprotonation enthalpy,

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

Table 1. Amino Acids Studied in This Work and Their Experimental pK_a Data at 298 K

				exp pK_a (298 K)		
no.	amino acid	abbrev.	type[a]	pK_{a1}	pK_{a2}	pK_{a3}
1	glycine	Gly	α, A	2.35[b]	9.77[c]	
2	β-alanine	β-ala	β, A	3.55[d]	10.33[b]	
3	taurine	Tau	α	1.5[e]	9.06[b]	
4	sarcosine	Sar	A	2.21[f]	10.21[b]	
5	methionine	Meth	α, B	2.13[g]	9.3[b]	
6	proline	Pro	α, A	1.95[h]	10.76[b]	
7	6-aminohexanoicacid	6-AHA	α	4.37[i]	10.80[i]	
8	phenylalanine	Phe-ala	α, B	2.2[j]	9.31[j]	
9	glutamic acid	Glu	α, A	2.19[k]	10.1[k]	4.45[k]
10	aspartic acid	Asp	α, A	1.99[i]	10.002[i]	3.9[i]

[a]A: Nonessential amino acid. B: Essential amino acid. [b]King et al.[121] [c]Hamborg et al.[122] [d]May et al.[123] [e]Andrews et al.[124] [f]Datta et al.[125] [g]Pelletier et al.[126] [h]Smith et al.[127] [i]Smith et al.[128] [j]Anderson et al.[129] [k]Nagai et al.[130]

Table 2. Gas Phase Basicities (GB)[a] of Data Set of Amino Acids Studied in This Work at 298 K. (All values are in kJ mol^{-1}.)

amino acid	G3MP2B3	G3MP2	G4MP2	CBS-QB3	DFT(B3LYP/6-311++G(d,p))	exp.
Gly	857	860	858	857	859	848[b], 850[c], 854[d]
β-ala	1099	1097	1098	1096	1093	
Tau	887	891	889	880	883	
Sar	887	892	887	884	888	882[e], 888.7[f]
Meth	898	900	900	894	897	892[b], 898[g], 899[d]
Pro	908	911	909	905	908	897[b], 904[h], 908[d]
6-AHA	967	945	967	964	966	
Phe-ala	895	896	895	894	894	887[b], 892[d]
Glu	916	919	916	913	914	902[b], 904[i], 908[d]
Asp	890	891	890	888	900	882[d], 875[f]

[a]The gas phase basicity is calculated using eq 3. [b]Harrison et al.[114] [c]Meot-Ner et al.[59] [d]Bouchoux et al.[110] [e]Locke et al.[58] [f]Hunter et al.[64] [g]Desaphy et al.[113] [h]Kuntz et al.[116] [i]Bouchoux et al.[111]

Table 3. Gas Phase Proton Affinities (PA)[a] of Data Set of Amino Acids Studied in This Work at 298 K. (All values are in kJ mol^{-1}.)

			G4MP2					
amino acid	G3MP2B3	G3MP2	this work	Gronert[b]	Uddin[c]	CBS-QB3	DFT(B3LYP/6-311++G(d,p))	exp.
Gly	892	889	892	887	890	887	885	881[d], 883[e], 887[g]
β-ala	1131	1133	1131			1128	1125	
Tau	916	918	916			910	920	
Sar	921	921	922			918	917	915[i], 921.2[j]
Meth	927	926	928	936	937	924	932	938[d,f,g]
Pro	938	940	940	942	939	936	943	929[d], 937[k], 942[g]
6-AHA	997	986	997			995	999	
Phe-ala	927	930	927	926	924	926	931	920[d], 931[l], 930[g]
Glu	948	948	949	948	948	946	943	947[d,g], 945[h]
Asp	923	923	924	916	917	921	918	905[d], 920[h], 908.9[j]

[a]The gas phase basicity is calculated using eq 4. [b]Gronert et al.[131] [c]Uddin et al.[76] [d]Harrison et al.[114] [e]Meot-Ner et al.[59] [f]Desaphy et al.[113] [g]Bouchoux et al.[110] [h]Bouchoux et al.[111] [i]Locke et al.[58] [j]Hunter et al.[64] [k]Kuntz et al.[116] [l]Bouchoux et al.[112]

and protonation entropy) of amino acids are also studied with composite and density functional methods. Calculated thermochemical properties are also compared with available literature data at 298 K. In the literature, gas phase thermochemistry of amino acids has been investigated by using equilibrium method,[57−60] the bracketing method,[61−64] and the kinetic method.[65−67] Gas phase thermochemical properties have also been studied computationally by using different methods.[68−73] Experimental gas phase IR spectra have been compared with computed IR spectra using density functional theory (DFT) by Oomens et al.[74] Recently Bouchoux et al.[75] and Uddin et al.[76] studied thermochemical properties of amino acids using DFT and composite methods. Stover et al.[77] provided a detailed discussion of the acidities of amino acids at the G3MP2 level plus some additional higher level calculations. Gas-phase acidities of aspartic acid, glutamic acid, and their amino acid amides have been studied by Li et al.[78] with density functional and molecular orbital theory approaches. In this work, we aim to provide a comprehensive study of thermochemical properties of a set of amino acids and compare with available experimental data.

2. COMPUTATIONAL DETAILS

In this work, for gaseous phase calculations, initial conformer search was carried out at the HF/6-31G* level and the most stable conformers at this level of theory for gas phase geometries of protonated, neutral, and deprotonated amino acids were fully

Table 4. Gas Phase Acidity (ΔG^0_{acid})[a] of Data Set of Amino Acids Studied in This Work at 298 K. (All values are in kJ mol^{-1}.)

amino acid	G3MP2B3	G3MP2	G4MP2	CBS-QB3	DFT(B3LYP/6-311++G(d,p))	exp.
Gly	1399	1400	1399	1392	1387	1404[b], 1400[c]
β-ala	1201	1202	1202	1198	1205	
Tau	1286	1283	1286	1287	1178	
Sar	1398	1397	1393	1392	1388	1399[b], 1400[c]
Meth	1374	1373	1373	1372	1366	1376[b]
Pro	1396	1394	1396	1391	1390	1395[b]
6-AHA	1392	1414	1392	1390	1375	
Phe-ala	1384	1384	1386	1380	1379	1379[b]
Glu	1340	1336	1341	1335	1312	1331[b]
Asp	1337	1336	1338	1332	1289	

[a]The gas phase acidity (ΔG^0_{acid}) is calculated using eq 5. [b]O'Hair et al.[117] [c]Locke et al.[58]

Table 5. Gas Phase Deprotonation Enthalpies (ΔH^0_{acid})[a] of Data Set of Amino Acids Studied in This Work at 298 K. (All values are in kJ mol^{-1}.)

amino acid	G3MP2B3	G3MP2	G4MP2	CBS-QB3	DFT(B3LYP/6-311++G(d,p))	exp.
Gly	1431	1433	1431	1428	1428	1434[b], 1433[c]
β-ala	1235	1231	1235	1231	1231	
Tau	1321	1318	1324	1322	1322	
Sar	1428	1429	1428	1425	1425	1429[c]
Meth	1412	1412	1411	1408	1408	1407[b], 1405[d]
Pro	1428	1427	1428	1423	1423	1431[b], 1430[d]
6-AHA	1437	1448	1436	1435	1435	
Phe-ala	1414	1421	1415	1410	1410	1418[b], 1408[d]
Glu	1368	1368	1369	1363	1358	1348[b], 1350[e]
Asp	1365	1365	1366	1360	1356	1345[b]

[a]The Gas phase Deprotonation Enthalpies (ΔH^0_{acid}) is calculated using eq 6. [b]Jones et al.[115] [c]Locke et al.[58] [d]O'Hair et al.[117] [e]Fournier et al.[132]

Table 6. Calculated Gas Phase[a] ΔS^0 of data set of amino acids studied in this work at 298 K. (All values are in J mol^{-1} K^{-1}.)

	G3MP2B3		G3MP2					
amino acid	this work	Uddin et al.	this work	Uddin et al.	G4MP2	CBS-QB3	DFT(B3LYP/6-311++G(d,p))	HF/6-31G*
Gly	−115.6	−109.6	−96.5	−122.7	−115.9	−101.9	−88.6	−99.3
β-ala	−108.9		−121.4		−110.0	−109.4	−108.5	−122.8
Tau	−96.6		−91.7		−91.9	−98.7	−124.2	−95.7
Sar	−115.5		−94.3		−116.5	−116.4	−96.8	−97.1
Meth	−96.1	−126.6	−89.5	−116.1	−97.2	−101.7	−117.3	−98.4
Pro	−103.0	−103.0	−96.6	−103.0	−103.8	−102.5	−119.5	−99.0
6-AHA	−101.6		−137.6		−103.4	−104.1	−108.1	−98.3
Phe-ala	−107.4	−99.7	−114.1	−95.5	−109.5	−107.3	−125.3	−98.5
Glu	−109.3	−130.3	−97.5	−124.2	−113.5	−111.9	−97.6	−103.0
Asp	−111.9	−114.1	−107.5	−107.9	−114.9	−112.9	−107.2	−108.7

[a]The gas phase ΔS^0 is calculated using eq 7.

optimized at the B3LYP level of theory using the 6-311++G (d, p) basis set. The most stable conformer was used for further thermodynamic calculations. Gas phase optimization of protonated, neutral, and deprotonated amino acids were carried out in Spartan 08. Frequency calculations confirmed absence of any imaginary frequencies in the minima. Total enthalpies and free energies in gas phase were obtained using Gaussian −n theories (G3MP2B3, G3MP2, G4MP2, CBS-QB3) and density functional theories (DFT) at B3LYP/6-311++G (d,p) level. Calculations for G3MP2B3, G3MP2, CBS-QB3, and DFT level of theories were done in Gaussian 03.[79] G4MP2 thermochemical calculations were done in Gaussian 09.[80]

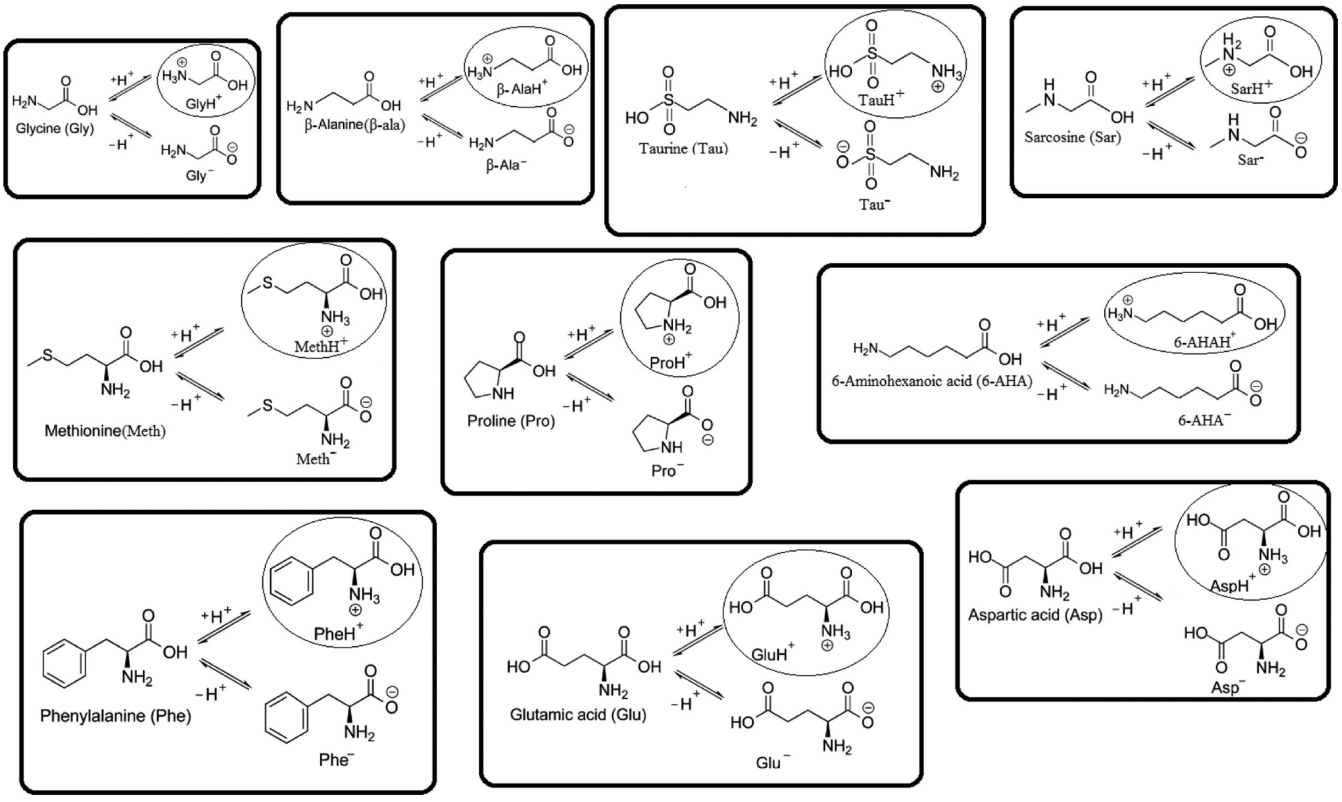
Aqueous phase geometries were fully optimized at B3LYP level of theory using 6-311++G (d, p) functional in Spartan 08 for PCM and SM8T calculations using solvent model SM8. PCM calculations were done using the default settings in Gaussian 03 in aqueous phase. All calculations were done using Density Functional Theory (DFT) at PCM/B3LYP/6-311++G (d, p)//SM8/B3LYP-6-311++G (d, p) level. The implementation of the PCM model[81] can be invoked using the Self-Consistent Reaction Field (SCRF) keyword in combination with PCM-specific modifiers. Option read was used to indicate a separate section of options providing calculation parameters for specifying the characteristics of the cavity. The extra options used were RADII = UAHF, which uses the United Atom Topological Model applied on radii optimized for the HF/6-31G (d) level of theory. The calculations done in this work use cavities based on atomic spheres, using the GEPOL algorithm of Nilsson et al.[82] GEPOL has been adopted as the default option for PCM calculations in the Gaussian 03 electronic structure program. It gives geometry optimization and energy calculations with better results. Both electrostatic and nonelectrostatic

Table 7. Comparison of Calculated Gas Phase Protonation Entropies ($\Delta_p S^0$)[a] with Computational Results from Literature and Experimental Data for Data Set of Amino Acids Studied in This Work at 298 K. (All values are in J mol^{-1} K^{-1}.)

	G3MP2B3		G3MP2					
amino acid	this work	Uddin et al.	this work	Uddin et al.	G4MP2	CBS-QB3	DFT(B3LYP/6-311++G(d,p))	exp.
Gly	−6.8	−0.8	12.3	−13.9	−7.1	6.9	20.2	−0.7, −1.9
β-ala	−0.1		−12.6		−1.2	−0.6	0.3	
Tau	12.2		17.1		16.9	10.1	−15.4	
Sar	−6.7		14.5		−7.7	−7.6	12.0	
Meth	12.7	−17.8	19.3	−7.3	11.6	7.1	−8.5	42.8, −25.4
Pro	5.8	5.8	12.2	5.8	5.0	6.3	−10.7	−0.7, −1.9
6-AHA	7.2		−28.8		5.4	4.7	0.7	
Phe-ala	1.4	9.1	−5.3	13.3	−0.7	1.5	−16.5	−0.7, −22.1
Glu	−0.5	−21.5	11.3	−15.4	−4.7	−3.1	11.2	−42.8, −28.8
Asp	−3.1	−5.3	1.3	0.9	−6.1	−4.1	1.6	−0.7

[a]The gas phase protonation entropies $\Delta_p S^0$ are calculated using eq 7.

Scheme 2. Gas Phase Protonation and Deprotonation Reaction Schematics of Amino Acids Studied in Present Work[a]



[a]The circle indicates the preferred protonation site.

(i.e., cavitation, repulsion, and dispersion) terms were included in the calculation of ΔG_{solv} values.

SM8T calculations[83] were done in Gamessplus,[84] using the optimized structure obtained earlier. All SM8T calculations were also done using Density Functional Theory at SM8T/B3LYP/6-311++G (d, p)//SM8/B3LYP-6-311++G (d, p) level. The SM8T model is not parametrized for ionic molecules but is expected to give good qualitative results for ions.[83]

The solvation shell geometries for the explicit solvation shell model were extracted from molecular simulations of the solute in the bulk solvent, and these clusters were fully optimized with quantum mechanical calculations as explained by da Silva et al.[51] The continuum models employed for calculations of cluster solvation energies was the Poisson−Boltzmann-based model in the DivCon code.[85] The continuum solvation calculations were carried out as single point calculations on the optimized HF/6-31+G (d) clusters. The Poisson−Boltzmann model calculations were carried out at the AM1 level (the model is not implemented at HF level), as explained in da Silva et al.[51] DivCon model results for single solutes were also taken into the present work for comparison purposes. All quantum mechanical calculations were carried out in Gaussian 03 software and all simulations were carried out using Sander from the AMBER 12 suite.[86] Details of the molecular dynamics simulations are given in the Supporting Information.

3. THEORETICAL BACKGROUND AND METHODS

3.1. Gas Phase Thermochemical Properties of Amino Acids. Gas phase thermochemical quantities associated with

basicity and acidity of an amino acid molecule (M) are defined as standard enthalpy and standard free energy of reactions 1 and 2, at 298 K,[87,88] as shown in Scheme 1.

$$MH_2^+ \rightarrow MH + H^+ \quad PA(MH) = \Delta H_1^0 \quad \text{and} \quad GB(M) = \Delta G_1^0 \tag{1}$$

$$MH \rightarrow M^- + H^+ \quad \Delta H_{acid}^0(MH) = \Delta H_2^0 \quad \text{and} \quad \Delta G_{acid}^0 = \Delta G_2^0 \tag{2}$$

Starting from these definitions, gas phase basicity (GB) and proton affinity (PA) for amino acids may be calculated from eq 1.

$$GB = [\Delta G^0(MH) - \Delta G^0(MH_2^+)] + \Delta G^0(H^+) \tag{3}$$

$$PA = [\Delta H^0(MH) - \Delta H^0(MH_2^+)] + \Delta H^0(H^+) \tag{4}$$

where $\Delta G^0(H^+) = -26.28$ kJ mol^{-1} and $\Delta H^0(H^+) = 6.19$ kJ mol^{-1} at 298 K.

Similarly, acidity ΔG_{acid}^0 and deprotonation enthalpy $\Delta H_{acid}^0(MH)$ of amino acids can be calculated from eq 5 and 6, following eq 2, as follows

$$\Delta G_{acid}^0 = [\Delta G^0(M^-) - \Delta G^0(MH)] + \Delta G^0(H^+) \tag{5}$$

$$\Delta H_{acid}^0 = [\Delta H^0(M^-) - \Delta H^0(MH)] + \Delta H^0(H^+) \tag{6}$$

Gas phase (ΔS^0) and "protonation entropy" terms have been introduced in the JANAF tables,[89] which relates to gas phase basicity and proton affinity as follows

$$GB = PA - T\Delta S^0 = PA - T[S^0(H^+) - \Delta_p S^0(MH)] \tag{7}$$

where $\Delta_p S^0 = S^0(MH_2^+) - S^0(MH)$ and $S^0(H^+) = 108.8$ J mol^{-1} K^{-1} at 298 K, calculated using Sackur–Tetrode equation.[90] This definition of protonation entropy may be extended to the deprotonation reaction, eq 2, as follows:

$$\Delta G_{acid}^0 = \Delta H_{acid}^0 + T\Delta S^0 = \Delta H_{acid}^0 + T[S^0(H^+) - \Delta_p S^0(M^-)] \tag{8}$$

3.2. Thermodynamic Cycle for Explicit Solvation Shell Model Solvation Energy Calculation. Cluster/continuum calculations of solvation free energies of ions with different thermodynamic cycles have been reported by several research groups.[30,31,37−40,43,47−49] The thermodynamic cycle for computing hydration free energies with cluster-continuum models has been discussed in detail[50,51] and is only briefly summarized here. The thermodynamic cycle used in the present work is given in Figure 1.

The upper leg of the water cluster cycle (Figure 1) involves gas phase reactions between the solute, A, and clusters of water molecules (referred to as $(H_2O)_n$).

The solvation free energies, $\Delta G_{solv}^*(A)$ of the solute were calculated via the thermodynamic cycle shown in Figure 1 as

$$\Delta G_{solv}^*(A) = \Delta G_{clust,g}^*(A(H_2O)_n) + \Delta G_{solv}^*(A(H_2O)_n) - \Delta G_{solv}^*((H_2O)_n) - \Delta G^{0\rightarrow *} - \Delta G^{*\rightarrow l} \tag{9}$$

This is the sum of the free energy of forming the gas phase solute–water cluster ($\Delta G_{clust,g}^*(A(H_2O)_n)$) with n explicit water molecules and the difference between the hydration free energies of the solute–water cluster ($\Delta G_{solv}^*(A(H_2O)_n)$) and the water cluster, $\Delta G_{solv}^*((H_2O)_n)$. The standard state corrections adjust the gas phase concentrations ($\Delta G^{0\rightarrow *} = RT \ln(24.46)$) from 1 mol per 24.46 L to 1 M and the water cluster concentration from 1 M to $55.34/n$ M[91,92] ($\Delta G^{*\rightarrow l} = RT \ln([H_2O]/n)$). The gas-phase standard state correction ($\Delta G^{0\rightarrow *}$) is 1.89 kcal mol^{-1} at room temperature. The study of standard state corrections for calculation of free energy of solvation can also be found in Gutowski et al.[93] The above cycle was used to study the solvation

Table 8. Free Energy of Solvation of Amino Acids Calculated by Explicit Solvation Shell Model (ESS). (All values are in kcal mol^{-1}.)

amino acid species (charge)	ΔG_{solv} (calcd)[a]	$\Delta G_{cluster}^*$[b]	$-T\Delta S_{cluster}^*$[c]	$\Delta G_s(A(S)_n)$[d]	area[e]
A (Cation)					
Gly (+1)	−78.30	−39.89	9.89	−60.66	217.08
β-ala (+1)	−70.84	−33.32	9.45	−59.33	231.90
Tau (+1)	−82.08	−37.86	11.02	−67.59	241.76
Sar (+1)	−67.22	−33.57	11.39	−57.39	230.04
Meth (+1)	−65.60	−30.41	10.46	−58.01	279.32
Pro (+1)	−61.06	−30.03	10.58	−53.97	247.18
6-AHA (+1)	−56.81	−26.87	10.30	−52.59	275.67
Phe-ala (+1)	−67.78	−35.31	9.63	−54.46	296.23
Glu (+1)	−67.06	−28.56	9.57	−60.44	272.92
Asp (+1)	−73.48	−33.42	9.75	−62.18	257.15
B (Zwitterion)					
Gly (0)	−31.86	−31.32	12.92	−32.82	202.64
β-ala (0)	−59.54	−45.46	13.64	−40.07	216.03
Tau (0)	−42.39	−31.06	12.62	−36.30	224.89
Sar (0)	−32.30	−28.13	12.73	−29.25	219.53
Meth (0)	−28.50	−21.33	12.20	−31.72	268.97
Pro (0)	−29.72	−27.66	12.50	−26.91	236.04
6-AHA (0)	−27.75	−27.27	13.36	−26.19	253.18
Phe-ala(0)	−37.69	−32.59	12.33	−29.79	285.27
Glu (0)	−31.13	−20.69	12.22	−35.01	259.05
Asp (0)	−35.1	−24.60	12.53	−35.42	246.94
C (Anion)					
Gly (−1)	−74.22	−32.03	10.09	−64.65	214.11
β-ala (−1)	−69.13	−30.80	12.14	−62.82	227.27
Tau (−1)	−71.46	−27.81	11.00	−67.02	241.32
Sar (−1)	−73.32	−35.79	10.56	−60.45	229.44
Meth (−1)	−67.56	−29.46	11.04	−61.51	282.03
Pro (−1)	−70.18	−34.77	11.57	−59.34	246.71
6-AHA (−1)	−72.59	−28.45	11.92	−68.41	283.15
Phe-ala (−1)	−68.90	−32.79	11.86	−60.33	292.83
Glu (−1)	−73.67	−28.93	12.30	−71.42	265.81
Asp (−1)	−79.77	−30.24	13.11	−74.99	247.02
D (Dianion)					
Glu (−2)	−201.19	−54.15	12.29	−171.69	274.04
Asp (−2)	−207.75	−56.41	12.83	−176.53	252.43

[a]Calculated free energy of solvation; all values shifted by −2.41 kcal/mol to remove systematic error relative to experimental values as in ESS model presented by da Silva et al.[51] Estimated sampling standard deviation is 1 kcal/mol. [b]Energy of formation of the cluster at the HF/6-31+G (d) level, converted from a standard state of 1 atm to 1 mol/L. Thermal corrections to the energy and zero-point energies not included. [c]Temperature (298 K) multiplied by the entropy of formation of the cluster at the HF/6-31+G (d) level. [d]Free energy of solvation of the cluster calculated with the Poisson–Boltzmann continuum model. [e]Area of clusters calculated with the Poisson–Boltzmann continuum model.

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

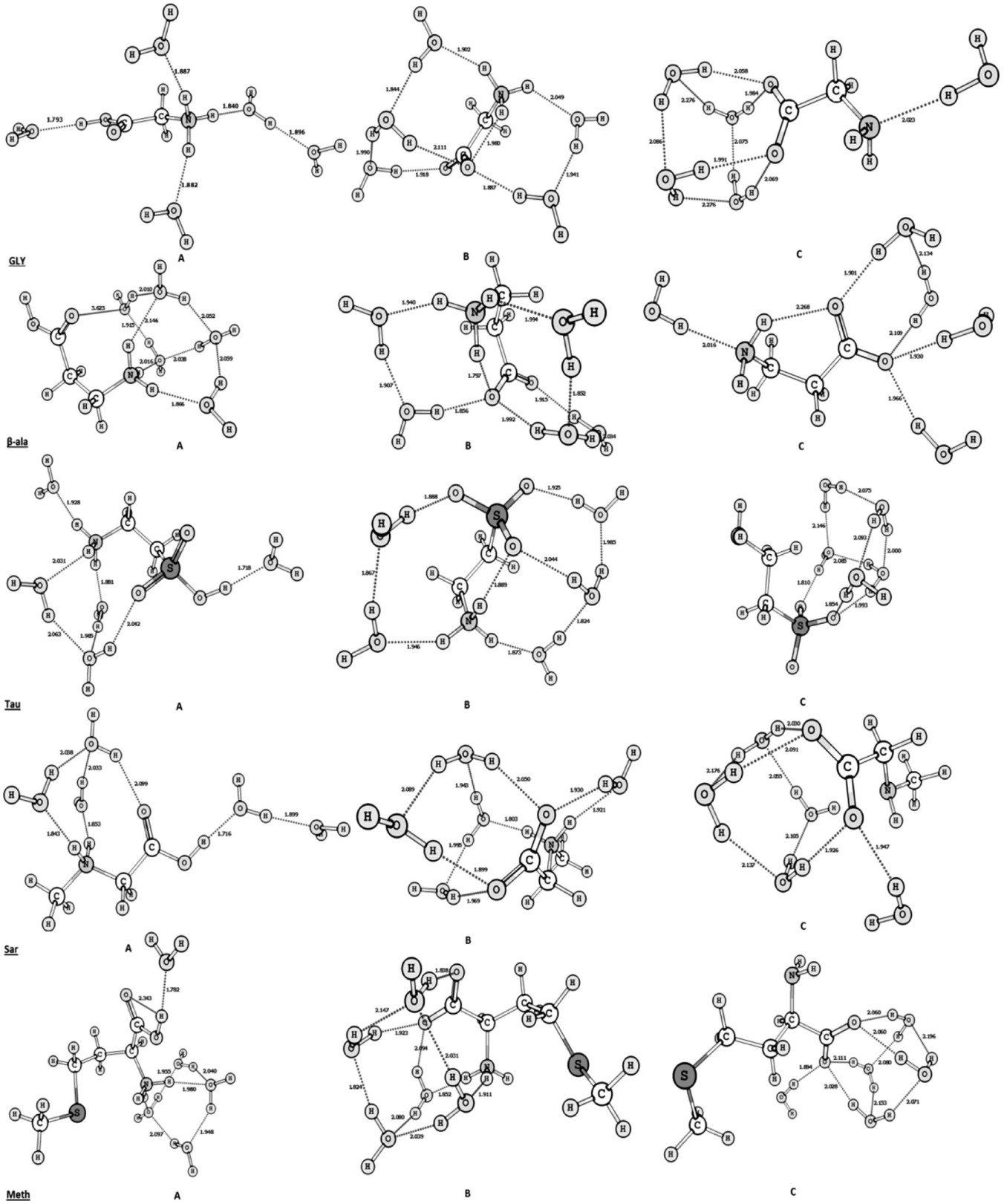


Figure 8. Optimized clusters of amino acids (Gly, β-ala, Tau, Sar, Meth) studied in this work. (A, B, C forms are explained in Figure 2). (Dotted lines shows hydrogen bonds, and bond lengths of hydrogen bonds are given in Angstrom.)

of a large series of cations and anions with clusters containing five explicit water molecules.[51] More recently, the similar cycle was used for the reference hydrated metal cation state in a study of metal binding specificity for the copper efflux regulator (CueR),[94] which used clusters of six water molecules for each metal ion: Ag^+, Cu^+, Au^+, Zn^{2+}, and Hg^{2+}, and by Riccardi et al.[95] to calculate hydration free energies of anions and group 12 divalent cations.

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

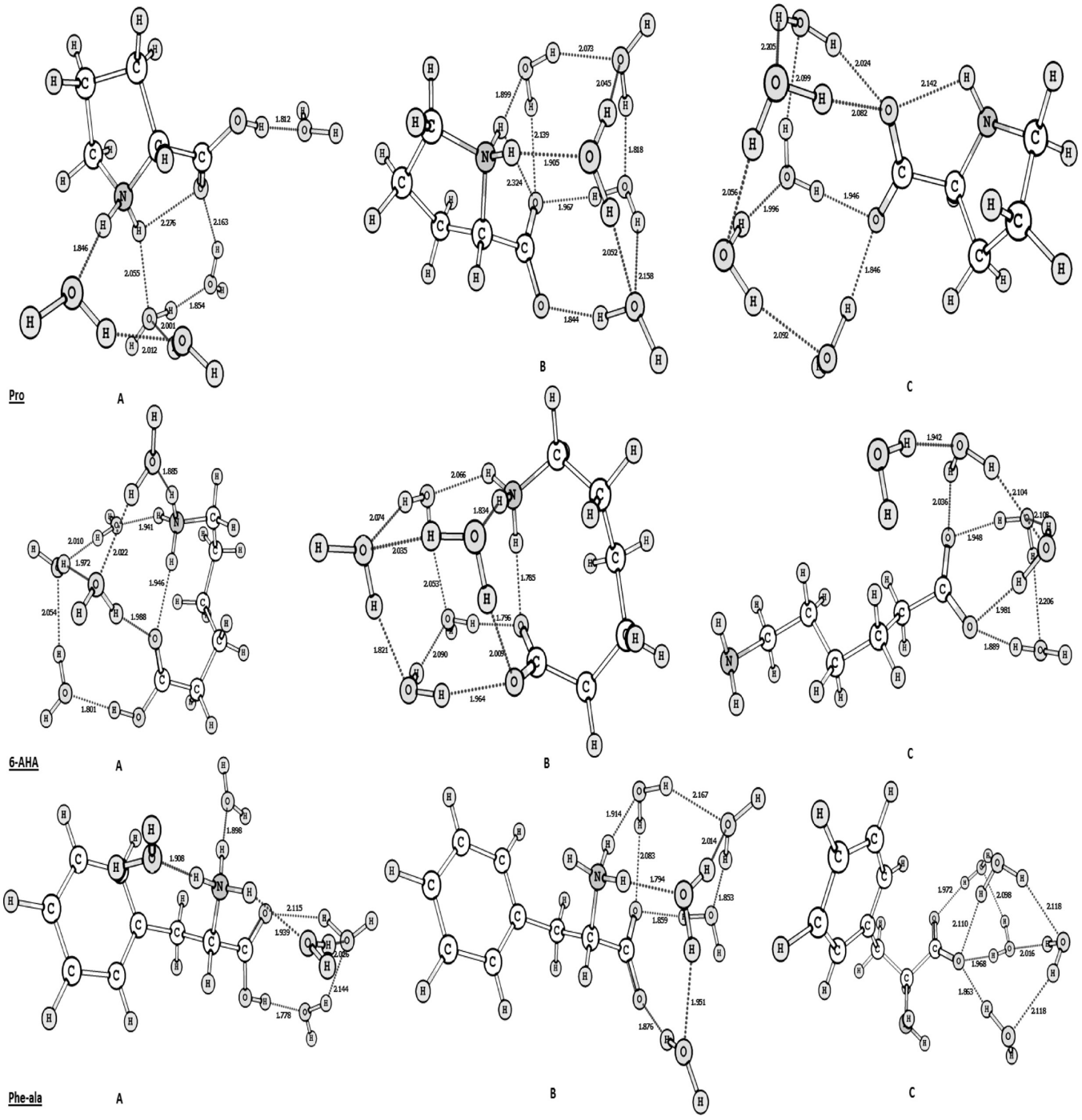


Figure 9. Optimized clusters of amino acids (Pro, 6-AHA, Phe-ala) studied in this work. (A, B, C forms are explained in Figure 2). (Dotted lines shows hydrogen bonds, and bond lengths of hydrogen bonds is given in Angstrom.)

3.3. pK_a Determination. The ionization constants, K_1 and K_2, which correspond to the reactions 1 and 2, respectively, are represented by

$$^{+}NH_3RCOOH + H_2O \rightarrow H_3O^{+} + {}^{+}NH_3RCOO^{-} \quad (10)$$

$$^{+}NH_3RCOO^{-} + H_2O \rightarrow H_3O^{+} + NH_2RCOO^{-} \quad (11)$$

Figure 2 shows the different species involved in amino acid pK_a determination. In case the amino acid has a side chain with an acid or basic group, the removal of the side chain acidic or basic hydrogen is termed pK_{a3}. pK_a determination of amino acids having acidic or basic side chains involve different species shown in Figure 3.

In the present study, we have two amino acids that have side chain carboxyl group (glutamic acid and aspartic acid).

pK_{a1}, pK_{a2}, and pK_{a3} values for amino acids calculated in the present work were obtained by using corresponding thermodynamic cycles shown in Figures 4−6.

Free energy of protonation in solution (ΔG^{*}_{aq}) for any deprotonation reaction, for example,

$$AH^{+}_{S} \xrightarrow{\Delta G^{*}_{aq,deprot(AH^{+})}} A_S + H^{+}_{S} \quad (12)$$

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

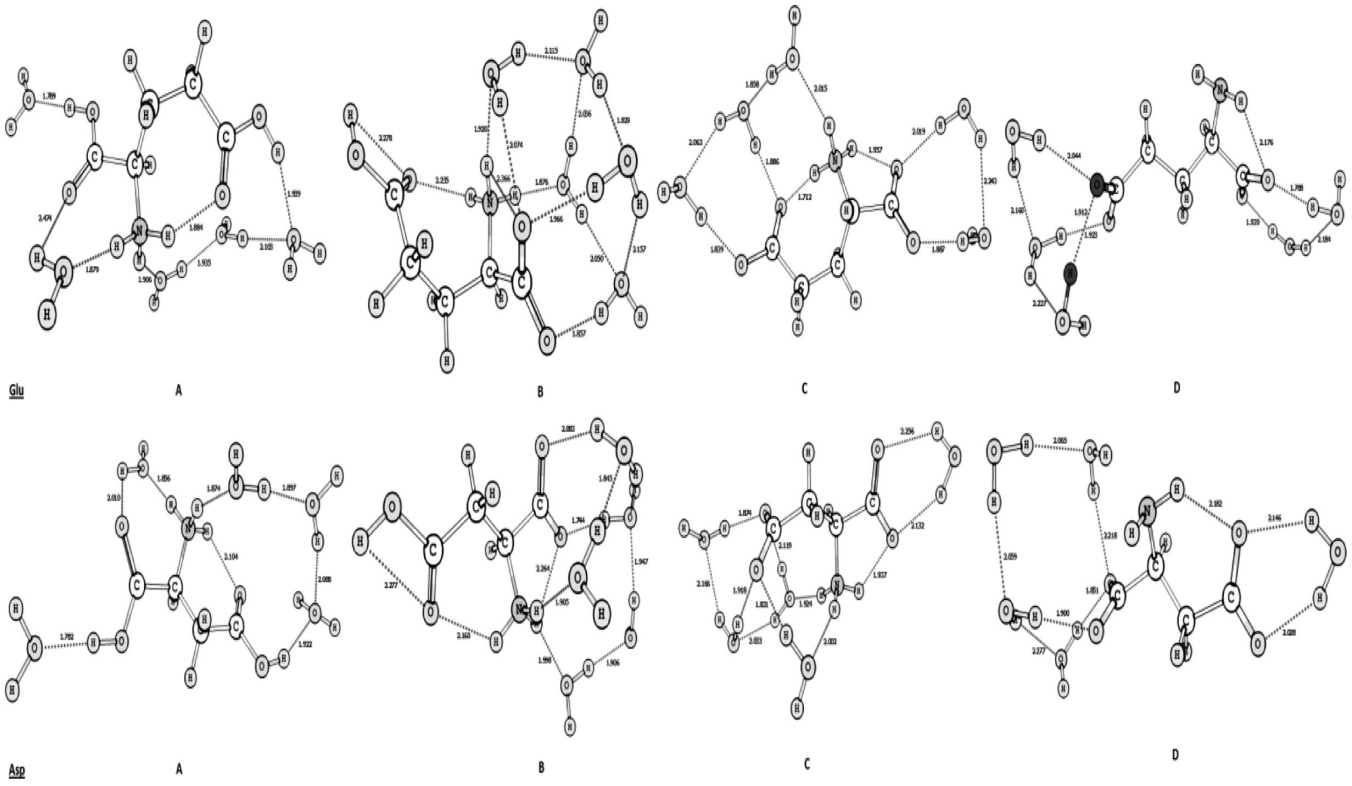


Figure 10. Optimized clusters of amino acids (Glu, Asp) studied in this work. (A, B, C, D forms are explained in Figure 2). (Dotted lines shows hydrogen bonds, and bond lengths of hydrogen bonds is given in Angstrom.)

is related to the aqueous pK_a according to

$$pK_a = \frac{\Delta G^*_{aq}}{RT\ \ln(10)} \tag{13}$$

The full method describing the calculation of free energy of protonation in solution as a summation of aqueous and gaseous solvation free energies is described in our previous study.[96]

3.4. Proton Exchange between Zwitterionic and Neutral Forms of Amino Acids. The potential energy surface of amino acids in the gas phase has been extensively studied by various researchers.[97−99] Zwitterionic forms of amino acids in vacuum are not found to be the minimum in ab initio calculations on the PES calculations using large basis sets.[100] Spectroscopic studies have also confirmed more stability of neutral forms of most of amino acids in gas phase.[101,102] However, various studies in literature shows that the zwitterionic form of most of amino acids dominates in crystalline or aqueous media.[103−108] In the present study, we have found that the zwitterionic form is stable in gas phase only at SCF level, where it gives a minimum at HF/6-31G* level in PES. This is shown in Scheme 1. So, zwitterion gas phase calculations in the present work were carried out at HF/6-31G* level and single point B3LYP/6-311++G(d,p)//HF/6-31G*. For calculating gas phase contribution in pK_a determination, energy difference between neutral and zwitterionic form of amino acid at single point DFT level is added to calculated CBS-QB3 gas phase energies for neutral form of amino acid. The energy difference in neutral and zwitterionic form of amino acids studied at HF/6-31G* level and single point B3LYP/6-311++G(d,p)//HF/6-31G* is given in Supporting Information (Table S2). Out of 10 amino acids studied in this work, β-alanine and taurine exist in their zwitterionic form at all composite and DFT methods studied in this work. So, no correction was added for β-alanine and taurine molecules. Figure 7 shows most stable conformer of amino acid (zwitterion or neutral) at CBS-QB3 level, taken for pK_a determination in this work.

3.5. Cluster Configurations. da Silva et al.[51] discuss the motivation for employing five explicit water molecules in an explicit solvation model. We have retained five explicit water molecules in this work also because of three underlying motivations. First, instead of adding a different number of water molecules for different molecules, remaining consistent with five water molecules for each molecule helps in cancelation of errors coming from gas phase free energy, entropy, and continuum solvation energy calculations, as they would be the same for each molecule. Second, to maintain computational cost of calculations within reasonable limits is also one concern. Third, the amino acid molecules studied in the present work are small organic molecules that can be reasonably solvated explicitly with five water molecules. Also, results presented by da Silva[51] were shown to be encouraging enough to proceed with the present studies of amino acids with the same number of water molecules, although there are many views on the issue of number of solvent molecules to be added to exactly represent solute−solvent interactions. Recent studies[47−50] have shown that methods such as adding solvent molecules until the calculated solvation energies converge could be more reliable than methods that are based on size and polarization of ionic molecules. However, these methods would be computationally expensive and would possibly not result in complete cancelation of error within one studied data set of molecules.

4. RESULTS

The data set of 10 amino acids studied in the present work is given in Table 1. This data set of amino acids is as stated above selected based on their potential for being solvents for Post Combustion CO_2 Capture Processes.[109]

The results for gas phase basicity, proton affinity, ΔG^0_{acid}, ΔH^0_{acid}, protonation entropy for amino acids studied in this work at different levels of theory are given in Tables 2−7.

Schematics of protonation and deprotonation reactions of amino acids studied in this work are shown in Scheme 2. From the results in Tables 2−7 for gas phase thermochemical properties, we can see that all composite methods and DFT method calculate gas phase thermochemical properties of amino acids, which agree fairly well with experimental results. The error bars in different theories are discussed in detail in discussion section.

Table 8 shows the calculated solvation energies obtained by employing the explicit solvation shell model presented by da Silva et al.[51] The cluster formation energies, entropies and cluster solvation energies for a set of 10 amino acids studied in this work are also given in Table 8. In this table results are presented for cationic, zwitterionic, anionic, and dianionic forms of the amino acid as explained in Figures 2 and 3. The cluster solvation energies in Table 8 are calculated with the Poisson−Boltzmann continuum solvation model. All calculations and corrections to the final results were done as explained in the model presented by da Silva et al.[51]

As in da Silva et al.,[51] in the present work, we also observed some molecule geometries that could not attain a minima and have imaginary frequencies, which resulted in breakdown of the geometry of the solute or solvent molecules. Vibration frequency calculations also had imaginary frequencies in some cases. Poor initial geometries obtained from molecular dynamics simulations can be the reason for such geometries. In order to remove such geometries and to retain results only from stable cluster geometries, we assigned an energy cut off of ±80 kcal/mol from the obtained minimum energy structure. Out of 100 geometries submitted for each molecule we obtained approximately 10−15 geometries for each molecule that could not converge to stable structure. Figures 8−10 show the most stable clusters obtained in the present work for species A, B, C, and D explained in Figure 2 and Figure 3 of the amino acids studied in present work.

Table 9 lists the free energy of solvation of each species of amino acid (A, B, C, D; i.e., cation, zwitterion, anion, and dianion) calculated with ESS. As discussed earlier, the ESS model is developed for predicting solvation free energies of ions. In Table 9, free energy of solvation of ionic amino acid species calculated from the continuum solvation shell model viz. the Polarized continuum solvation model (PCM), SM8T, and DivCon are also given.

Results for pK_{a1}, pK_{a2}, and pK_{a3} calculated by the Explicit solvation shell model, the PCM method, the SM8T method, and DivCon method are listed in Tables 10, 11, and 12, respectively, along with the respective experimental values. All results are presented at 298 K. Unsigned mean errors in pK_a relative to experimental pK_a from various models for every amino acid are also given in these tables.

In order to analyze each leg of the relevant thermodynamic cycle in detail to determine how different terms contribute to the observed unsigned mean error (UME) of pK_a between theory and experiment. pK_a based on free energy of solution values obtained from PCM and SM8T optimization calculations at PCM/B3LYP/6-311++G(d,p) and SM8T/B3LYP/6-311++G(d,p) levels of theory respectively is determined. The results for pK_{a1}, pK_{a2}, and pK_{a3} calculated from this approach are given in Supporting Information (Table S3−S5). No significant improvement in results was observed by following lower leg of relative thermodynamic cycle for determination of pK_a values.

Table 9. Comparison of Free Energy of Solvation Calculated by Using Explicit Solvation Shell (ESS) Model and Implicit Solvation Models (PCM, SM8T, DivCon). (All values are in kcal mol^{-1}.)

	ΔG_{solv} (calcd)			
	explicit solvation shell model	implicit solvation shell models		
amino acid species (charge)	ΔG_{solv} (calcd)[a]	PCM	SM8T	DivCon
	A (Cation)			
Gly (+1)	−78.30	−81.72	−86.33	−81.17
β-ala (+1)	−70.84	−73.04	−76.15	−74.34
Tau (+1)	−82.08	−78.32	−86.18	−87.83
Sar (+1)	−67.22	−76.38	−75.15	−70.3
Meth (+1)	−65.60	−66.15	−69.65	−70.24
Pro (+1)	−61.06	−71.05	−68.72	−64.83
6-AHA (+1)	−56.81	−59.4	−61.46	−61.8
Phe-ala (+1)	−67.78	−69.96	−72.87	−70.23
Glu (+1)	−67.06	−74.85	−73.99	−70.55
Asp (+1)	−73.48	−81.39	−79.12	−76.06
	B (Zwitterion)			
Gly (0)	−31.86	−31.56	−33.18	−33.95
β-ala (0)	−59.54	−54.35	−62.31	−60.21
Tau (0)	−42.39	−31.13	−41.82	−41.32
Sar (0)	−32.30	−29.88	−28.96	−30.38
Meth (0)	−28.50	−24.82	−26.32	−30.60
Pro (0)	−29.72	−29.51	−27.50	−28.51
6-AHA (0)	−27.75	−19.83	−21.86	−26.03
Phe-ala(0)	−37.69	−28.44	−30.63	−34.47
Glu (0)	−31.13	−35.7	−34.95	−32.60
Asp (0)	−35.13	−38.18	−36.61	−35.62
	C (Anion)			
Gly (−1)	−74.22	−71.30	−73.63	−74.26
β-ala (−1)	−69.13	−73.04	−74.93	−72.53
Tau (−1)	−71.46	−68.54	−78.33	−74.27
Sar (−1)	−73.32	−70.27	−73.505	−72.73
Meth (−1)	−67.56	−63.80	−65.17	−69.66
Pro (−1)	−70.18	−69.63	−72.25	−71.19
6-AHA (−1)	−72.59	−75.46	−76.77	−74.67
Phe-ala (−1)	−68.90	−64.67	−65.76	−70.2
Glu (−1)	−73.67	−73.65	−70.53	−77.27
Asp (−1)	−79.77	−76.99	−74.21	−81.41
	D (Dianion)			
Glu (−2)	−201.19	−190.63	−198.47	−200.75
Asp (−2)	−207.75	−203.27	−215.80	−210.16

[a]Calculated free energy of solvation; all values shifted by −2.41 kcal/mol to remove systematic error relative to experimental values as in ESS model presented by da Silva et al.[51] Estimated sampling standard deviation is 1 kcal/mol.

5. DISCUSSION

5.1. Comparison of Gas Phase Thermochemistry of Amino Acids from Different Levels of Theory and Experimental Results. The gas phase basicity, gas phase proton affinities, gas phase acidity, and gas phase deprotonation enthalpy for amino acids are given in Table 2−5. The difference between experimental and calculated gas phase thermochemical properties (basicity (GB), gas phase proton affinity (PA), gas phase acidity (ΔG^0_{acid}) and gas phase deprotonation enthalpy (ΔH^0_{acid}) are plotted in Figure 11. There have been many studies in the literature for gas phase thermochemistry of amino acids.[64,76,87,110−118] The main source of experimental values for

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

Table 10. Comparison of Calculated pK_{a1} from Explicit Solvation Shell Model and Implicit Solvation Shell Models with Experimental pK_{a1} at 298 K

		pK_{a1} at 298 K				unsigned mean error (UME)			
amino acid	exp.	ESS	PCM	SM8T	DivCon	ESS	PCM	SM8T	DivCon
Gly	2.35[a]	7.97	15.85	18.05	13.69	5.62	13.50	15.70	11.34
β-ala	3.55[b]	8.14	13.58	10.01	10.22	4.59	10.03	6.46	6.67
Tau	1.5[c]	4.62	10.14	8.06	9.64	3.12	8.64	6.56	8.14
Sar	2.21[d]	−1.28	7.25	7.02	2.40	3.49	5.04	4.81	0.19
Meth	2.13[e]	1.04	4.15	5.62	2.90	1.09	2.02	3.49	0.78
Pro	1.95[f]	2.32	9.82	9.59	5.98	0.36	7.87	7.64	4.03
6-AHA	4.37[g]	5.34	13.08	13.11	10.29	0.97	8.71	8.73	5.91
Phe-ala	2.2[h]	6.38	14.79	15.31	10.55	4.18	12.59	13.11	8.35
Glu	2.19[i]	6.41	8.78	8.70	7.89	4.22	6.59	6.51	5.70
Asp	1.99[g]	−2.17	1.41	0.90	−0.63	4.16	0.58	1.10	2.62
average error						3.18	7.56	7.41	5.38

[a]King et al.[121] [b]May et al.[123] [c]Andrews et al.[124] [d]Datta et al.[125] [e]Pelletier et al.[126] [f]Smith et al.[127] [g]Smith et al.[128] [h]Anderson et al.[129] [i]Nagai et al.[130]

Table 11. Comparison of Calculated pK_{a2} from Explicit Solvation Shell Model and Implicit Solvation Shell Models with Experimental pK_{a2} at 298 K

		pK_{a2} at 298 K				unsigned mean error (UME)			
amino acid	exp.	ESS	PCM	SM8T	DivCon	ESS	PCM	SM8T	DivCon
Gly	9.77[a]	11.49	13.45	12.93	13.03	1.72	3.68	3.16	3.26
β-ala	10.33[b]	10.70	4.01	8.47	8.70	0.37	6.32	1.86	1.63
Tau	9.06[b]	12.13	5.99	6.65	9.27	3.07	3.07	2.41	0.21
Sar	10.21[b]	11.24	11.71	8.64	10.27	1.03	1.50	1.57	0.06
Meth	9.3[b]	10.21	10.28	10.37	10.22	0.91	0.98	1.07	0.92
Pro	10.76[b]	18.37	18.62	15.21	16.74	7.61	7.86	4.45	5.98
6-AHA	10.80[c]	11.27	3.33	3.85	8.47	0.47	7.48	6.95	2.33
Phe-ala	9.31[d]	8.92	12.54	13.35	12.91	0.39	3.23	4.04	3.60
Glu	10.1[e]	11.41	19.17	11.10	14.38	1.30	9.06	0.99	4.28
Asp	10.002[c]	10.91	12.17	0.90	10.35	0.90	2.16	9.11	0.34
Average Error						1.78	4.53	3.56	2.26

[a]Hamborg et al.[122] [b]King et al.[121] [c]Smith et al.[128] [d]Anderson et al.[129] [e]Nagai et al.[130]

Table 12. Compariosn of Calculated pK_{a3} from Explicit Solvation Shell Model and Implicit Solvation Shell Models with Experimental pK_{a3} at 298 K

		pK_{a3} at 298 K				unsigned mean error (UME)			
amino acid	exp	ESS	PCM	SM8T	DivCon	ESS	PCM	SM8T	DivCon
Glu	4.45[a]	2.90	6.28	8.03	1.33	1.55	1.83	3.58	3.12
Asp	3.9[b]	2.15	6.45	7.33	1.31	1.75	2.55	3.43	2.59
average error						1.65	2.19	3.50	2.85

[a]Nagai et al.[130] [b]Smith et al.[128]

GB and PA is reviewed in a paper by Harrison et al.[114] They reported average values for GB and PA derived from various literature sources. Experimental sources from literature for different thermochemical property of amino acids are reported in Tables 2−7.

From Figure 11, we can see that the difference between the results for gas phase thermochemical properties given by B3LYP/6-311++G (d,p) level of theory and Gaussian −n theories and the experimental results for amino acids are within ±10 kJ/mol. Only for gas phase deprotonation enthalpy of glutamic acid and aspartic acid are these errors as large as ±20 kJ/mol. Gaussian-n theory based calculations are within ±5 kJ/mol experimental error bars in most of molecules for gas phase thermochemical properties. Different experimental sources of data can vary by ∼2% relative to each other.[76] Different Gaussian-n theory results (G3MP2, G3MP2B3, G4MP2, and CBS-QB3) in this work agree with each other to within ∼1% or less in terms of relative difference (5 kJ mol^{-1}). So, it can be suggested that these differences are within the range of experimental error bars and can likely be considered to be random in nature.

The gas phase (ΔS^0) and protonation entropy ($\Delta_p S^0$) calculated from GB and PA using eq 7 are given in Tables 6 and 7, respectively. Calculated ΔS^0 and $\Delta_p S^0$ were compared with results from G3MP2B3 and G3MP2 levels of theory from Uddin et al.,[76] and both agree within ±10 kJ/mol.

5.2. Comparison of Continuum Solvation Models and Explicit Solvation Shell Model Free Energy of Solvation. In Tables 8 and 9, we have given the free energy of solvation results calculated by using the explicit solvation shell model presented by da Silva et al.[51] and continuum solvation models (PCM, SM8T, and DivCon). By comparing the results calculated

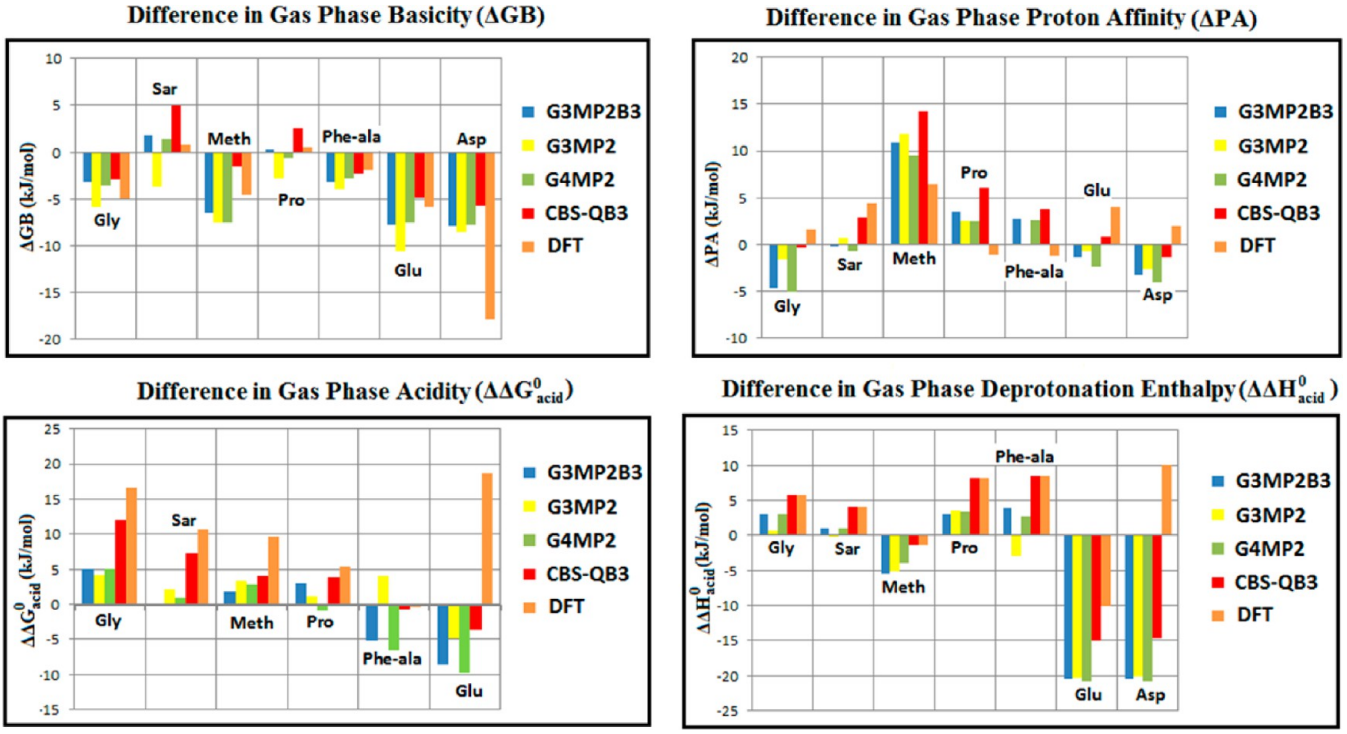


Figure 11. Plot of G3MP2B3, G3MP2, G4MP2, CBS-QB3, DFT (B3LYP/6-311++G (d,p)) gas phase basicity difference ($\Delta GB = GB_{exp} - GB_{calcd}$), gas phase proton affinity difference ($\Delta PA = PA_{exp} - PA_{calcd}$), gas phase acidity difference ($\Delta\Delta G^0_{acid} = \Delta G^0_{acid(exp)} - \Delta G^0_{acid(calcd)}$), gas phase deprotonation enthalpy difference ($\Delta\Delta H^0_{acid} = \Delta H^0_{acid(exp)} - \Delta H^0_{acid(calcd)}$) in kJ mol^{-1} amino acids studied in this work.

Table 13. Comparison of Different Absolute Errors for pK_{a1}, pK_{a2}, and pK_{a3} from Explicit Solvation Shell Model and Continuum Solvation Models (PCM, SM8T, DivCon)

absolute errors[a]	ESS	PCM	SM8T	DivCon
		pK_{a1}		
SME	−1.43	−7.44	−7.19	−4.85
UME	3.18	7.56	7.41	5.38
RMSE	3.59	8.53	8.47	6.33
		pK_{a2}		
SME	−1.85	−0.05	0.01	−1.26
UME	1.78	4.53	3.56	2.26
RMSE	2.74	5.30	4.37	2.95
		pK_{a3}		
SME	1.65	−2.19	−3.50	2.85
UME	1.65	2.19	3.50	2.85
RMSE	1.65	2.22	3.50	2.86

[a]Different errors are abbreviated as SME (signed mean error), UME (unsigned mean error), and RMSE (root mean square error).

Table 14. Comparison of Different Relative Errors for pK_{a1}, pK_{a2}, and pK_{a3} from Explicit Solvation Shell Model and Continuum Solvation Models (PCM, SM8T, DivCon)

relative errors	ESS	PCM	SM8T	DivCon
		pK_{a1}		
SME	0.00	0.00	0.00	0.00
UME	2.82	3.34	3.29	3.40
RMSE	3.30	4.16	4.47	4.06
		pK_{a2}		
SME	0.00	0.00	0.00	0.00
UME	1.54	4.51	3.56	2.21
RMSE	2.16	5.29	4.37	2.56
		pK_{a3}		
SME	0.00	0.00	0.00	0.00
UME	0.65	0.36	0.07	0.26
RMSE	0.10	0.36	0.07	0.26

with explicit solvation shell model with that of continuum solvation models, we cannot make any absolute judgment, as experimental ΔG_{solv} values are not available for the amino acids studied in the present work. However, looking at results given by ESS, PCM, SM8T, and DivCon continuum solvation models, we can see that all these models predicts amino acid zwitterion free energy of solvation in same range, though we know that these models have different levels of theoretical treatment and parametrization for molecules. We would like to emphasis that all models in the study have some degree of empirical fitting, but none of them have been fitted to any of the molecules in the present study and no parameter fitting is done in the present work.

5.3. Comparison of Performance of ESS Model and Continuum Solvation Models for pK_a Calculation.

Comparison of absolute and relative pK_a errors from ESS and continuum solvation models for pK_{a1}, pK_{a2}, and pK_{a3}, respectively, is given in Table 13 and 14. Absolute errors are errors calculated by subtracting calculated pK_a from corresponding experimental pK_a. Relative errors are errors obtained by first adding the signed mean error (SME) from every model to their corresponding calculated pK_a results and then subtracting this resulting calculated pK_a from experimental pK_a values.

Absolute Errors. Table 13 lists absolute errors in pK_{a1}, pK_{a2}, and pK_{a3} calculated from the Explicit Solvation Shell and continuum solvation models (PCM, SM8T, DivCon). Comparing the experimental pK_a from the *Handbook of Chemistry and Physics* (2003) and other previous values available in the

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

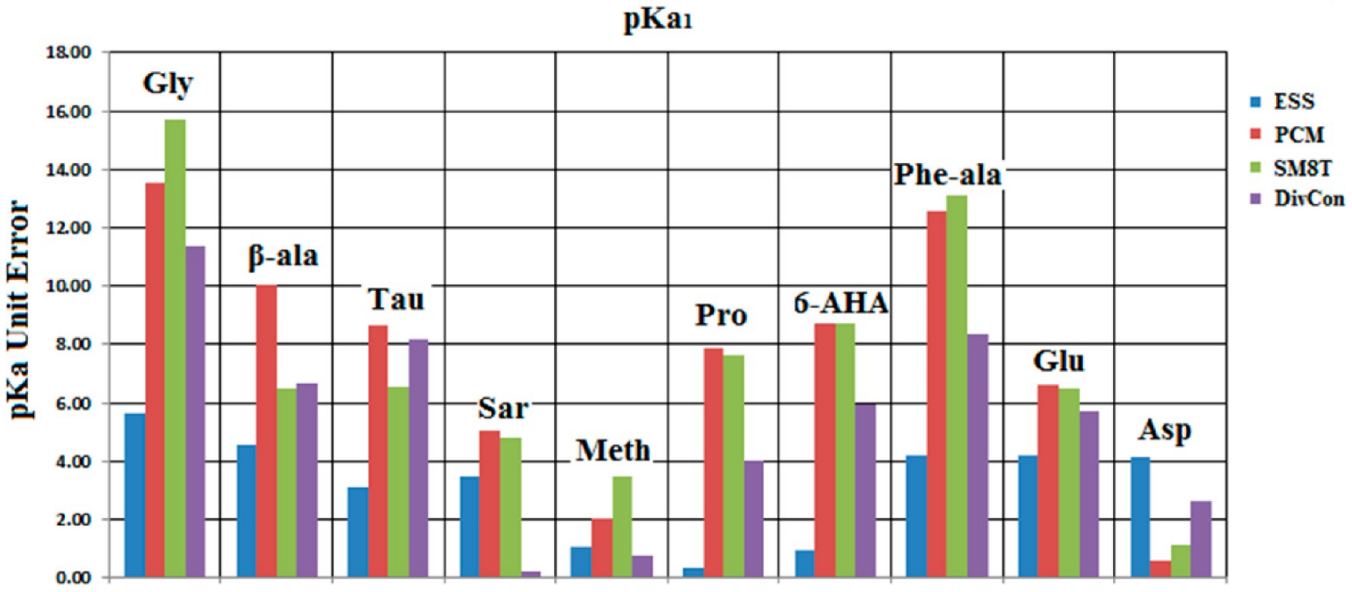


Figure 12. Graphical representation of pK_a unit errors in the calculation of pK_{a1} of amino acids using explicit solvation shell model (ESS) and continuum solvation models (PCM, SM8T, and DivCon).

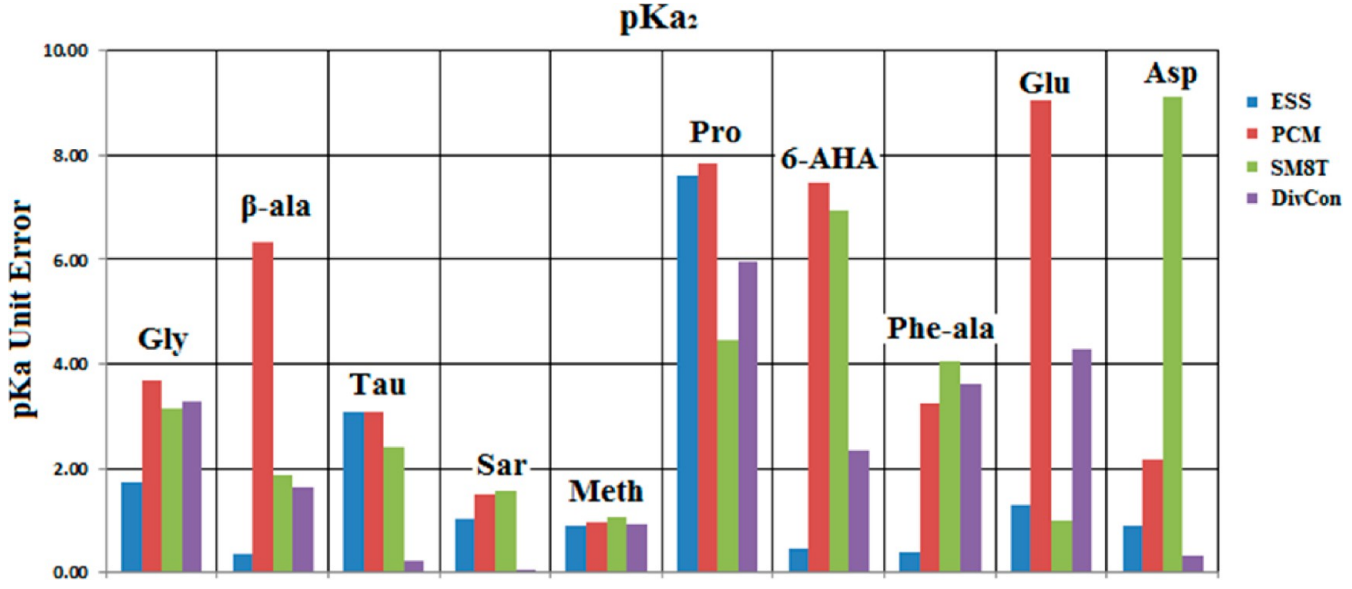


Figure 13. Graphical representation of pK_a unit errors in the calculation of pK_{a2} of amino acids using explicit solvation shell model (ESS) and continuum solvation models (PCM, SM8T, and DivCon).

literature given in Table 1 also show slight differences. This comparison of experimental pK_a values for current set of amino acids is also given in our recent study.[119]

By comparing absolute signed mean error (SME), unsigned mean error (UME) and root-mean-square errors (RMSE) for the set of 10 amino acid studied in this work, the much higher accuracy of absolute pK_a prediction from the ESS model over continuum solvation models is easily visible. The ESS model gives a UME error of 3.18, 1.78, and 1.65 pK_a units in pK_{a1}, pK_{a2}, and pK_{a3} respectively. The ESS model gives better results than all continuum solvation models. We are getting an absolute mean unsigned error (UME) of 7.56, 4.53, and 2.19 pK_a units in pK_{a1}, pK_{a2}, and pK_{a3}, respectively, with PCM; 7.41, 3.56, and 3.50 pK_a units in pK_{a1}, pK_{a2}, and pK_{a3}, respectively, with SM8T; and 5.38, 2.26, and 2.85 pK_a units in pK_{a1}, pK_{a2}, and pK_{a3}, respectively, with DivCon solvation model. The absolute signed mean errors (SME) and unsigned mean errors (UME) in pK_{a1}, pK_{a2}, and pK_{a3} given by ESS are very low in comparison to the corresponding errors from continuum solvation models so the stability of the explicit solvation shell model for prediction of free energy of ions and pK_a values is found to be very good. Mangold et al.[120] reports unsigned mean error of 2.1 pK_a units with a maximum error of 4.0 pK_a units in their density functional based molecular dynamics simulation study of data set of 6 amino acids. The root mean squared errors (RMSE) of 3.30, 2.16, and 1.65 in pK_{a1}, pK_{a2}, and pK_{a3} given by ESS are very encouraging when considering the fact that the ESS model does not have any parametrization for any molecule in the present study.

dx.doi.org/10.1021/ct400459y | J. Chem. Theory Comput. 2013, 9, 5021−5037

Absolute errors in amino acid pK_{a1} and pK_{a2} with ESS and continuum solvation models (PCM, SM8T, and DivCon) are also plotted in Figures 12 and 13. From Figures 12 and 13, it can easily be seen that explicit solvation shell model gives much less errors in pK_a units in comparisons to continuum solvation models.

Relative Errors. Based on the definition of relative error, as explained above, the relative SME will be zero for every model. The ESS model gives less relative UME and RMSE compared to the PCM, SM8T, and DivCon continuum solvation models. Relative RMSE given by ESS for pK_{a1}, pK_{a2}, and pK_{a3} is 3.30, 2.16, and 0.10 pK_a units respectively. A comparison of relative errors given by ESS and continuum solvation models of the present study is given in Table 14.

5.4. Free Energy of Solvation of Zwitterions ($\Delta G_{solv}^{Zwitterions}$). No computational model has been parametrized for zwitterions as there is no experimental data available. The reason is that it is difficult to study intermediate molecules. In Table 9, we list free energies of solvation of the amino acid zwitterions ($\Delta G_{solv}^{Zwitterions}$) studied in this work calculated from ESS and continuum solvation models (PCM, SM8T, and DivCon). By comparing the results we can see that all models predicts $\Delta G_{solv}^{Zwitterions}$ in the same range, and it becomes interesting when we understand that the various methods used for the calculations in ESS, PCM, SM8T, and DivCon are substantially different in theory from each other. As discussed earlier, we believe the explicit solvation shell model approach is one of the best methods available today for studying the free energy of solvation of ionic molecules. Thus, we believe that the $\Delta G_{solv}^{Zwitterions}$ predicted with the ESS model in this work are reasonable predictions for these values.

6. CONCLUSIONS

The absolute and relative errors in pK_a values of amino acids predicted from the explicit solvation model are lower than for continuum models studied in this work. Due to experimental difficulties in determining solvation free energies of amino acids, solvation free energies given in the present work for a set of amino acids can serve as guiding values for further development in this area. We hope that our results will encourage theoretical chemists and experimentalists to develop new ways to determine solvation free energies of amino acids. Highly precise solvation free energies of amino acids not only might serve to validate our data but may also be invaluable for the refinement of simulation methods and theoretical understanding of the solvation contributions to all bimolecular processes. This work also provides a consistent set of gas phase thermochemical properties that can serve as benchmarks. Gas phase thermochemical properties calculated with composite methods in this work are found to be consistent with available experimental data. For amino acid molecules for which experimental data are not available, calculated gas phase thermochemical properties can serve as good estimates.

■ ASSOCIATED CONTENT

Supporting Information

Constant terms utilized in calculations for ESS model. Underlying data for Table 2, energy of solute and cluster in gas phase, thermal corrections to the energy and entropy of solute and cluster. Gaseous phase free energy for amino acids at 298 K using HF/6-31G* level. Simulation details. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: hallvard.svendsen@chemeng.ntnu.no

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Financial support for this work by Aker Clean Carbon, EON, EnBW, and the Norwegian Research Council CLIMIT program through the SOLVit project is greatly appreciated.

■ REFERENCES

(1) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2007**, *111*, 408.
(2) Marcus, Y. *Ion Solvation*; Wiley: Chichester, U.K., 1985.
(3) Abraham, M. H.; Andonianhaftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. *J. Chem. Soc. Perkin Trans. 2* **1994**, 1777.
(4) Betterton, E. A.; Hoffmann, M. R. *Environ. Sci. Technol.* **1988**, *22*, 1415.
(5) Cabani, S.; Conti, G.; Lepori, L. *Trans. Faraday Soc.* **1971**, *67*, 1933.
(6) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849.
(7) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. *J. Med. Chem.* **2008**, *51*, 769.
(8) Butler, J. A. V. *Trans. Faraday Soc.* **1937**, *33*, 229.
(9) Wolfenden, R. *J. Gen. Physiol.* **2007**, *129*, 357.
(10) Konig, G.; Bruckner, S.; Boresch, S. *Biophys. J.* **2013**, *104*, 453.
(11) Dill, K. A. *J. Biol. Chem.* **1997**, *272*, 701.
(12) Chen, J. H.; Brooks, C. L.; Khandogin, J. *Curr. Opin. Struc. Biol.* **2008**, *18*, 140.
(13) Feig, M.; Brooks, C. L. *Curr. Opin. Struc. Biol.* **2004**, *14*, 217.
(14) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
(15) Jeancharles, A.; Nicholls, A.; Sharp, K.; Honig, B.; Tempczyk, A.; Hendrickson, T. F.; Still, W. C. *J. Am. Chem. Soc.* **1991**, *113*, 1454.
(16) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.
(17) Guthrie, J. P. *Can. J. Chem.* **1991**, *69*, 1893.
(18) Guthrie, J. P. *Can. J. Chem.* **1992**, *70*, 1042.
(19) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292.
(20) Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 5665.
(21) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133.
(22) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *9*, 609.
(23) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517.
(24) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Aid. Mol. Des.* **1995**, *9*, 87.
(25) Klamt, A.; Schuurmann, G. *J. Chem. Soc. Perkin Trans. 2* **1993**, 799.
(26) Cramer, C. J.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 760.
(27) Camaioni, D. M.; Dupuis, M.; Bentley, J. *J. Phys. Chem. A* **2003**, *107*, 5778.
(28) Chipman, D. M. *J. Chem. Phys.* **2003**, *118*, 9937.
(29) Chipman, D. M.; Chen, F. W. *J. Chem. Phys.* **2006**, *124*, 144507.
(30) Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S. *J. Chem. Phys.* **2003**, *119*, 2702.
(31) Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B. *J. Am. Chem. Soc.* **2004**, *126*, 1285.
(32) De Abreu, H. A.; Guimaraes, L.; Duarte, H. A. *J. Phys. Chem. A* **2006**, *110*, 7713.
(33) Gutowski, K. E.; Dixon, D. A. *J. Phys. Chem. A* **2006**, *110*, 8840.
(34) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 16066.
(35) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 2493.

(36) Li, J.; Fisher, C. L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35*, 4694.
(37) Martin, R. L.; Hay, P. J.; Pratt, L. R. *J. Phys. Chem. A* **1998**, *102*, 3565.
(38) Mejias, J. A.; Lago, S. *J. Chem. Phys.* **2000**, *113*, 7306.
(39) Pliego, J. R.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241.
(40) Pliego, J. R.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *106*, 7434.
(41) Rode, B. M.; Schwenk, C. F.; Hofer, T. S.; Randolf, B. R. *Coordin. Chem. Rev.* **2005**, *249*, 2993.
(42) Rotzinger, F. P. *J. Am. Chem. Soc.* **1997**, *119*, 5230.
(43) Tawa, G. J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. *J. Chem. Phys.* **1998**, *109*, 4852.
(44) Uudsemaa, M.; Tamm, T. *J Phys Chem A* **2003**, *107*, 9997.
(45) Uudsemaa, M.; Tamm, T. *Chem. Phys. Lett.* **2004**, *400*, 54.
(46) Wiebke, J.; Moritz, A.; Cao, X.; Dolg, M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 459.
(47) Zhan, C. G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534.
(48) Zhan, C. G.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 9737.
(49) Zhan, C. G.; Dixon, D. A. *J. Phys. Chem. A* **2004**, *108*, 2020.
(50) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. *J. Phys. Chem. B* **2008**, *112*, 9709.
(51) da Silva, E. F.; Svendsen, H. F.; Merz, K. M. *J. Phys. Chem. A* **2009**, *113*, 6404.
(52) Grabowski, P.; Riccardi, D.; Gomez, M. A.; Asthagiri, D.; Pratt, L. R. *J. Phys. Chem. A* **2002**, *106*, 9145.
(53) Rempe, S. B.; Pratt, L. R.; Hummer, G.; Kress, J. D.; Martin, R. L.; Redondo, A. *J. Am. Chem. Soc.* **2000**, *122*, 966.
(54) Ho, J. M.; Coote, M. L. *J. Chem. Theory Comput.* **2009**, *5*, 295.
(55) Liptak, M. D.; Shields, G. C. *Int. J. Quantum Chem.* **2001**, *85*, 727.
(56) Marenich, A. V.; Ding, W. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. Lett.* **2012**, *3*, 1437.
(57) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695.
(58) Locke, M. J.; McIver, R. T. *J. Am. Chem. Soc.* **1983**, *105*, 4226.
(59) Meot-Ner, M.; Hunter, E. P.; Field, F. H. *J. Am. Chem. Soc.* **1979**, *101*, 686.
(60) Su, T.; Bowers, M. T. *J. Am. Chem. Soc.* **1973**, *95*, 7611.
(61) Carr, S. R.; Cassady, C. J. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1203.
(62) Cassady, C. J.; Carr, S. R.; Zhang, K.; Chungphillips, A. *J. Org. Chem.* **1995**, *60*, 1704.
(63) Gorman, G. S.; Speir, J. P.; Turner, C. A.; Amster, I. J. *J. Am. Chem. Soc.* **1992**, *114*, 3986.
(64) Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413.
(65) Bojesen, G.; Breindahl, T. *J. Chem. Soc. Perkin Trans. 2* **1994**, 1029.
(66) Isa, K.; Omote, T.; Amaya, M. *Org. Mass Spectrom.* **1990**, *25*, 620.
(67) Li, X. P.; Harrison, A. G. *Org. Mass Spectrom.* **1993**, *28*, 366.
(68) Alcami, M.; Mo, O.; Yanez, M. *Mass Spectrom. Rev.* **2001**, *20*, 195.
(69) Alcami, M.; Mo, O.; Yanez, M. *J. Phys. Org. Chem.* **2002**, *15*, 174.
(70) Bliznyuk, A. A.; Schaefer, H. F.; Amster, I. J. *J. Am. Chem. Soc.* **1993**, *115*, 5149.
(71) Li, Z.; Matus, M. H.; Velazquez, H. A.; Dixon, D. A.; Cassady, C. J. *Int. J. Mass Spectrom.* **2007**, *265*, 213.
(72) Tian, Z. X.; Pawlow, A.; Poutsma, J. C.; Kass, S. R. *J. Am. Chem. Soc.* **2007**, *129*, 5403.
(73) Topol, I. A.; Burt, S. K.; Russo, N.; Toscano, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 318.
(74) Oomens, J.; Steill, J. D.; Redlich, B. *J. Am. Chem. Soc.* **2009**, *131*, 4310.
(75) Bouchoux, G.; Huang, S. H.; Inda, B. S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 651.
(76) Uddin, K. M.; Warburton, P. L.; Poirier, R. A. *J. Phys. Chem. B* **2012**, *116*, 3220.
(77) Stover, M. L.; Jackson, V. E.; Matus, M. H.; Adams, M. A.; Cassady, C. J.; Dixon, D. A. *J. Phys. Chem. B* **2012**, *116*, 2905.
(78) Li, Z.; Matus, M. H.; Velazquez, H. A.; Dixon, D. A.; Cassady, C. J. *Int. J. Mass Spectrom.* **2007**, *265*, 213.
(79) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; 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 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(80) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
(81) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(82) Silla, E.; Villar, F.; Nilsson, O.; Pascualahuir, J. L.; Tapia, O. *J. Mol. Graphics* **1990**, *8*, 168.
(83) Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2008**, *112*, 3024.
(84) Higashi, M. M., A. V.; Olson, R. M.; Chamberlin, A. C.; Pu, J.; Kelly, C. P.; Thompson, J. D.; Xidos, J. D.; Li, J.; Zhu, T.; Hawkins, G. D.; Chuang, Y.-Y.; Fast, P. L.; Lynch, B. J.; Liotard, D. A.; Rinaldi, D.; Gao, J.; Cramer, C. J.; Truhlar, D. G. *GAMESSPLUS*, version 2010−2; University of Minnesota: Minneapolis, 2010, based on the General Atomic and Molecular Electronic Structure System (GAMESS) as described in Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347.
(85) Wang, B.; Raha, K.; Liao, N.; Peters, M. B.; Kim, H.; Westerhoff, L. M.; Wollacott, A. M.; van der Vaart, A.; Gogonea, V.; Suarez, D.; Dixon, S. L.; Vincent, J. J.; Brothers, E. N.; Merz, K. M. J. *DivCon*; QuantumBio: State College, PA, 2007.
(86) Case, D. A.; Darden, T. A.; Cheatham, T. E.III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; ; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A.W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12* ; University of California: San Francisco, 2012.
(87) Bouchoux, G. *Mass Spectrom. Rev.* **2007**, *26*, 775.
(88) Maksic, Z. B.; Kovacevic, B.; Vianello, R. *Chem. Rev.* **2012**, *112*, 5240.
(89) Chase, M. *NIST-JANAF Thermochemical Tables*; American Institute of Physics for the National Institute of Standards and Technology: [Washington, D.C.]; Woodbury, N.Y., 1998.
(90) McQuarrie, D. M. *Statistical Mechanics*; Harper and Row: New York, 1970; p 86
(91) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.
(92) Camaioni, D. M.; Schwerdtfeger, C. A. *J. Phys. Chem. A* **2005**, *109*, 10795.

(93) Gutowski, K. E.; Dixon, D. A. *J. Phys. Chem. A* **2006**, *110*, 8840.
(94) Rao, L.; Cui, Q. A.; Xu, X. *J. Am. Chem. Soc.* **2010**, *132*, 18092.
(95) Riccardi, D.; Guo, H.-B.; Parks, J. M.; Gu, B.; Liang, L.; Smith, J. C. *J. Chem. Theory Comput.* **2012**, *9*, 555.
(96) Gupta, M.; da Silva, E. F.; Svendsen, H. F. *J. Phys. Chem. B* **2012**, *116*, 1865.
(97) Hu, C. H.; Schaefer, H. F.; Hou, Z. L.; Bayes, K. D. *J. Am. Chem. Soc.* **1993**, *115*, 6904.
(98) Hu, C. H.; Shen, M. Z.; Schaefer, H. F. *J. Am. Chem. Soc.* **1993**, *115*, 2923.
(99) Császár, A. G. *J. Mol. Struct.* **1995**, *346*, 141.
(100) Ding, Y. B.; Kroghjespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261.
(101) Schafer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566.
(102) Godfrey, P. D.; Brown, R. D. *J. Am. Chem. Soc.* **1995**, *117*, 2019.
(103) Albrecht, G.; Corey, R. B. *J. Am. Chem. Soc.* **1939**, *61*, 1087.
(104) Jonsson, P. G.; Kvick, A. *Acta Crystallogr. B* **1972**, *B 28*, 1827.
(105) Kvick, A.; Liminga, R.; Jonsson, P. G. *Acta Chem. Scand.* **1972**, *26*, 1087.
(106) Tortonda, F. R.; PascualAhuir, J. L.; Silla, E.; Tunon, I. *Chem. Phys. Lett.* **1996**, *260*, 21.
(107) Leung, Y. C.; Marsh, R. E. *Acta Crystallogr.* **1958**, *11*, 17.
(108) Marsh, R. E. *Acta Crystallogr.* **1958**, *11*, 654.
(109) Hamborg, E. S.; Niederer, J. P. M.; Versteeg, G. F. *J. Chem. Eng. Data* **2007**, *52*, 2491.
(110) Bouchoux, G. *Mass Spectrom. Rev.* **2012**, *31*, 391.
(111) Bouchoux, G.; Bimbong, R. N. B.; Nacer, F. *J. Phys. Chem. A* **2009**, *113*, 6666.
(112) Bouchoux, G.; Bourcier, S.; Blanc, V.; Desaphy, S. *J. Phys. Chem. B* **2009**, *113*, 5549.
(113) Desaphy, S.; Malosse, C.; Bouchoux, G. *J. Mass Spectrom.* **2008**, *43*, 116.
(114) Harrison, A. G. *Mass Spectrom. Rev.* **1997**, *16*, 201.
(115) Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. *Int. J. Mass Spectrom.* **2007**, *267*, 54.
(116) Kuntz, A. F.; Boynton, A. W.; David, G. A.; Colyer, K. E.; Poutsma, J. C. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 72.
(117) Ohair, R. A. J.; Bowie, J. H.; Gronert, S. *Int. J. Mass Spectrom.* **1992**, *117*, 23.
(118) Stover, M. L.; Jackson, V. E.; Matus, M. H.; Adams, M. A.; Cassady, C. J.; Dixon, D. A. *J. Phys. Chem. B* **2012**, *116*, 2905.
(119) Gupta, M.; da Silva, E. F.; Svendsen, H. F. *J. Phys. Chem. B* **2013**, *117*, 7695.
(120) Mangold, M.; Rolland, L.; Costanzo, F.; Sprik, M.; Sulpizi, M.; Blumberger, J. *J. Chem. Theory Comput.* **2011**, *7*, 1951.
(121) King, E. J. *J. Am. Chem. Soc.* **1945**, *67*, 2178.
(122) Hamborg, E. S.; Niederer, J. P. M.; Versteeg, G. F. *J. Chem. Eng. Data* **2007**, *52*, 2491.
(123) May, M.; Felsing, W. A. *J. Am. Chem. Soc.* **1951**, *73*, 406.
(124) Andrews, S.; Schmidt, C. L. A. *J. Biol. Chem.* **1927**, *73*, 651.
(125) Datta, S. P.; Leberman, R.; Rabin, B. R. *Trans. Faraday Soc.* **1959**, *55*, 1982.
(126) Pelletier, S. Q., M. *Compt. Rend.* **1957**, *244*, 894.
(127) Smith, P. K.; Gorham, A. T.; Smith, E. R. B. *J. Biol. Chem.* **1942**, *144*, 737.
(128) Smith, E. R. B.; Smith, P. K. *J. Biol. Chem.* **1942**, *146*, 187.
(129) Anderson, K. P.; Greenhalgh, W. O.; Izatt, R. M. *Inorganic Chemistry* **1966**, *5*, 2106.
(130) Nagai, H.; Kuwabara, K.; Carta, G. *J. Chem. Eng. Data* **2008**, *53*, 619.
(131) Gronert, S.; Simpson, D. C.; Conner, K. M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2116.
(132) Fournier, F.; Afonso, C.; Fagin, A. E.; Gronert, S.; Tabet, J. C. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1887.