



HAL
open science

Hydration effects on the vibrational properties of carboxylates: From continuum models to QM/MM simulations

Vishal Kumar Porwal, Antoine Carof, Francesca Ingrosso

► **To cite this version:**

Vishal Kumar Porwal, Antoine Carof, Francesca Ingrosso. Hydration effects on the vibrational properties of carboxylates: From continuum models to QM/MM simulations. *Journal of Computational Chemistry*, 2023, 44 (23), pp.1898-1911. 10.1002/jcc.27171 . hal-04153406

HAL Id: hal-04153406

<https://hal.science/hal-04153406>

Submitted on 6 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hydration Effects on the Vibrational Properties of Carboxylates: From Continuum Models to QM/MM Simulations

Vishal Kumar Porwal¹ | Antoine Carof¹ | Francesca Ingresso¹

¹Université de Lorraine and CNRS,
Laboratoire de Physique et Chimie
Théoriques UMR 7019, Nancy, France

Correspondence

Email: Francesca.Ingresso@univ-lorraine.fr

Funding information

The presence of carboxyl groups in a molecule delivers an affinity to metal cations and a sensitivity to the chemical environment, especially for an environment that can give rise to intermolecular hydrogen bonds. Carboxylate groups can also induce intramolecular interactions, such as the formation of hydrogen bonds with donor groups, leading to an impact on the conformational space of biomolecules. In the latter case, the protonation state of the amino groups plays an important role. In order to provide an accurate description of the modifications induced in a carboxylated molecule by the formation of hydrogen bonds, one needs a compromise between a quantum chemical description of the system and the necessity to take into account explicit solvent molecules. In this work, we propose a bottom-up approach to study the conformational space and the carboxylate stretching band of (bio)organic anions. Starting from the anions in a continuum solvent, we then move to calculations using a microsolvation approach including one explicit water molecule per polar group, immersed in a con-

Abbreviations: LDH, layered double hydroxides; S symmetric and AS antisymmetric stretch, IR, infrared; DFT, Density functional theory; QM/MM, quantum mechanics/molecular mechanics; PES, potential energy surface; SMD, solvation model density; PCM, polarizable continuum model; H-bond, hydrogen bond; PMF, potential of mean force.

tinuum. Finally, we run QM/MM molecular dynamics simulations to analyze the solvation properties and to explore the anions conformational space. The results thus obtained are in good agreement with the description given by the microsolvation approach and they bring a more detailed description of the solvation shell and of the intermolecular hydrogen bonds.

KEYWORDS

Carboxylates, quantum chemistry, hydration effects, vibrational properties, QM/MM.

1 | INTRODUCTION

Molecules containing carboxylate groups are used in a wide range of applications. The structure as well as the electronic properties of the carboxylate moiety make it an ideal unit for the molecule to be exploited as a ligand. In coordination chemistry, carboxylates, in particular multi-carboxylates, behave as bridging ligands and are used to build coordination polymers and new materials.⁽¹⁾ They can be used to synthesize metal organic frameworks,⁽²⁾ photocatalytic materials,⁽³⁾ and inorganic open frame materials used for sorption and catalytic processes.⁽⁴⁾ Given their strong interaction with lanthanides, they are used to stabilize lanthanide ions in water solution and reduce their toxicity, thus allowing the development of luminescent probes for cell imaging.^(5,6)

In organic synthesis using atom and step economy principles, carboxylates can assist in the catalytic activity of Ru complexes.⁽⁷⁾ Polypyridine Ru complexes functionalized with carboxylate groups have been employed for the design of dye synthesized solar cells, since those group allow anchoring with the TiO₂ electrode surface, and in electrocatalysis.^(8,9)

The mechanism of carboxylate-metal in biological systems binding affects the affinity and selectivity of metals, and thus the function of metalloproteins.^(10,11) The affinity of carboxylates to metal cations is also exploited in the technology based on cellulose to produce nanofibrils, facilitating the process leading to biocompatible hydrogels,⁽¹²⁾ and in the mechanism leading to hydrogels based on alginates and pectin.^(13,14) In the field of biocompatible materials, the carboxylate units of amino acids, peptides and other biomolecules are involved as ligands in the formation of metal-biomolecule frameworks.⁽¹⁵⁾

The ability of carboxylates to act as anchoring groups is particularly important in the interaction with mineral surfaces. The possibility of mineral dissolution in water strongly depends on the interaction at the mineral surface/water interface. In principle, it is possible to control the solid growth or dissolution by triggering the adsorption of ions that are initially dispersed in the surrounding solution. In particular, polycarboxylates are used to control particle aggregation for calcite in water.⁽¹⁶⁾ In the field of metallurgy, the separation of minerals from ore bodies is often based on flotation through the use of surfactants containing fatty acids, which interact via coordination to calcium ions on the surface.⁽¹⁷⁾ Modifications of the wettability of reservoir rock, leading to an improved oil recovery process, can be induced though desorption of organic carboxylates from the rock surface by interaction with surfactants.⁽¹⁸⁻²⁰⁾

Organic and biological molecules with carboxylate groups can be adsorbed within the interlayer of claylike materials to build hybrid organic-inorganic systems. Layered double hydroxides (LDH) are anionic clays comprising M^{III}

58 and M^{II} cations, coordinated by six oxygen atoms belonging to hydroxyl groups, packing into layers. The surface
 59 charge is positive, so that the interlamellar space that hosts an aqueous environment intercalates anions of various
 60 nature, which leads to versatile and biocompatible materials.⁽²¹⁻²⁶⁾ LDH intercalating molecules with a negative charge
 61 born by carboxylate groups, such as deprotonated carboxylic acids, have been systematically synthesized since the
 62 1970's,⁽²⁷⁾ and it has been suggested that they can provide stabilization, transport and release of drugs, such as for
 63 instance ibuprofen, containing such groups.⁽²⁸⁾

64 Among the possible intercalates containing carboxylate groups, amino acids have attracted increasing attention
 65 for the production of bioinorganic materials.⁽²⁹⁻³¹⁾ Natural anionic clays might have played a role in concentrating
 66 amino acid units, providing a propitious environment for the polymerization leading to polypeptides, the first step
 67 towards the origin of life in abiotic conditions.⁽³²⁾ Aspartate, the deprotonated form of aspartic acid, has a deprotonated
 68 side chain at neutral and basic pH and it was present in the first naturally synthesized proteins.⁽³³⁾ In our group,
 69 we recently became interested in the mechanism of the anion-surface interaction, mediated by intercalated water
 70 molecules, and in the role played by the amino group of aspartate in anchoring the surface,⁽³⁴⁾ compared with anions
 71 lacking such unit, such as organic carboxylates. If we compare, for instance, the structure of succinate and aspartate,
 the only structural difference is the absence/presence of an amino group (see Scheme 1).

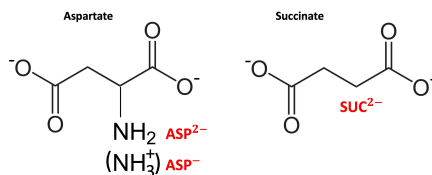


FIGURE 1 Structure of the aspartate anions (which will be named ASP^- and ASP^{2-} , depending on the protonation state) compared with the structure of the succinate anion (SUC^{2-}).

72

73 As the carboxylate group, the amino group has a polar character and can be involved in the formation of hydrogen
 74 bonds with water as well with the surface hydroxyl groups of LDH. Although remarkable insights were provided by
 75 applying molecular simulations on LDH intercalating carboxylates and amino acids,⁽³⁴⁻³⁸⁾ we are not aware of previous
 76 theoretical work explicitly examining how these different molecular structures can induce differences in the adsorption
 77 mechanism inside the material.

78 Experimentally, the carboxylate stretch band may be exploited to investigate the local structure surrounding
 79 the intercalates, as it is very sensitive to inter(intra)molecular interactions and provides a probe of the local environ-
 80 ment.^(17,39,40) Computational work has been carried out to elucidate hydration⁽⁴¹⁾, coordination⁽⁴²⁾ and both hydra-
 81 tion and coordination⁽⁴³⁾ effects on this infrared (IR) band for carboxylates. However, extensive computational studies
 82 on model systems are very rare in the literature. To our knowledge, a systematic study on how structural differences
 83 (e.g. succinate vs, aspartate), the protonation state of the amino group, the formation of intermolecular hydrogen
 84 bonds affects the carboxylate stretch band is still lacking, especially with respect to a comparison between different
 85 computational methods. As matter of fact, a fine description of the band position and shape requires sophisticated
 86 theoretical approaches (see for instance recent work by Vieira Pinto et al. 44 and references therein). On the other
 87 hand, a reasonable compromise can be achieved through normal mode analysis from quantum chemistry calculations.
 88 However, a validation of such an approach, possibly including a limited number of explicit water molecules, has to be
 89 provided on each considered system.⁽⁴¹⁾

90 As a first step in our study of molecules containing carboxylate groups, we focused on succinate (SUC^{2-}), zwitter-

91 ionic aspartate (ASP^-), the most stable form in a neutral water solution, and aspartate with a deprotonated amino
92 group (ASP^{2-}), the most stable form in basic solution and in the interlamellar space of LDH.⁽³⁷⁾ In this work, we concen-
93 trated our attention on applying different models to describe hydration effects and on comparing the results obtained
94 with different quantum chemistry methods. In a bottom-up approach, firstly we performed electronic structure cal-
95 culations using a continuum representation of the water solvent. We then included a number of water molecules
96 corresponding to the number of polar groups (two for succinate, three for aspartate),^(41,45) in a continuum and fi-
97 nally, once the PM6⁽⁴⁶⁾ semiempirical level was validated through comparison with density functional theory (DFT)
98 calculations, we ran hybrid quantum mechanics-molecular mechanics (QM/MM) molecular dynamics (MD). Such an
99 approach allowed us to investigate the local water network and the conformational space of the solute (by umbrella
100 sampling⁽⁴⁷⁾ along the C-C-C-C dihedral) in the presence of an explicit solvent. An evaluation of the IR spectrum in
101 the carboxyl stretch region was also provided by computing the power spectrum of the (QM) molecular dipole. We
102 would like to point out that the choice of including only a very limited number of water molecules in the quantum
103 calculations, instead of developing a more systematic microsolvation approach (for bioorganic molecules, see for in-
104 stance Refs. 48–55), was supported by the work of Sutton et al. on carboxylates⁽⁴¹⁾ and by Hernández et al. on
105 aspartate,⁽⁴⁵⁾ which provide an excellent framework for validation.

106 This paper is organized as follows. After a description of the computational procedures in Section 2, we describe
107 the results obtained based on electronic structure calculations (Section 3.1) and those based on QM/MM molecu-
108 lar dynamics (Section 3.2). Conclusions are drawn in Section 4, where some perspectives for future work are also
109 presented.

110 2 | COMPUTATIONAL DETAILS

111 We performed electronic structure calculations using Gaussian16,⁽⁵⁶⁾ and Gaussview6⁽⁵⁷⁾ for some analysis and for
112 creating images of the molecular structures. For each anion (SUC^{2-} , ASP^- and ASP^{2-} in Scheme 1), and for each of
113 the quantum chemistry methods that will be discussed in the following, we first performed a relaxed scan along the
114 C-C-C-C dihedral. We then performed a full optimization on the minima found along the constrained potential energy
115 surface (PES) and characterized them through an analysis of the frequencies of the normal modes (at 298 K and 1 bar).
116 The stationary points would be considered as minima only if no imaginary frequency was found. The normal mode
117 frequencies and intensities from the quantum calculations were processed by means of a convolution of Gaussian
118 functions having a full width at half maximum of 20 cm^{-1} , to obtain infrared bands.

119 The choice of the methods was based on recent work that has already been published in the literature: density
120 functional theory using the M05-2X⁽⁵⁸⁾ functional and the cc-pVDZ basis set,⁽⁵⁹⁾ used by Sutton et al. for carboxy-
121 lates among which succinate,⁽⁴¹⁾; DFT with the B3LYP⁽⁶⁰⁾ functional and the 6-311++G(d,p) basis set,^(61,62) used by
122 Hernández et al. for aspartate.⁽⁴⁵⁾ Results obtained with these methods have been compared with PM6⁽⁴⁶⁾ semiem-
123 pirical calculations. Two different continuum models of solvation were used, namely the solvation model density
124 (SMD)⁽⁶³⁾ and the polarizable continuum model (PCM).⁽⁶⁴⁾

125 For each minimum along the PES of the three molecules, we added one water molecule sitting next to the polar
126 groups (i.e. 2 water molecules around succinate - one per carboxylate group- and 3 water molecules around aspartate
127 - one per carboxylate groups and one for the amino group). The structures thus obtained were optimized and normal
128 modes analysis based on the harmonic approximation was performed to confirm the obtention of minima using the
129 same choice of methods already described for the anions, in the presence of water as a continuum medium. This pro-
130 cedure follows the ideas in Refs. 41,45, in which such microsolvation approach was applied to analyze the formation

of a solute-solvent hydrogen bond. In aspartate, such band is sensitive to couplings with the amino group,⁽⁶⁵⁾ and it is therefore important to assess the effect of the interaction of explicit water molecules with this group.

Having found a picture that is qualitatively the same for the different levels of theory, we proceeded to use PM6 based QM/MM molecular dynamics with the AMBER16⁽⁶⁶⁾ suite of programs. After minimization using the steepest descent algorithm, we performed a 10 ps equilibration in the NPT ensemble using Andersen's thermostat⁽⁶⁷⁾ (velocity rescaling every 1 ps) and Berendsen's barostat⁽⁶⁸⁾ in a cubic box containing one solute molecule and about 4000 water molecules, using periodic boundary conditions and a modified Ewald scheme for treating long range electrostatics in QM/MM based dynamics.⁽⁶⁹⁾ The force field used for water is TIP3P,⁽⁷⁰⁾ widely employed to describe water next to biological systems. The time step used in the simulations was 1 fs. After equilibration, each trajectory was propagated in NPT (1 bar, 300 K) for 500 ps and used for data analysis by employing post processing tools built in the AMBER16 package as well as in-house codes.

In order to have a qualitative description of the configurational space of the three anions in solutions, umbrella sampling⁽⁴⁷⁾ simulations were performed along the C-C-C-C dihedral (see Scheme 1) with the QM/MM scheme. The potential of mean force (PMF) was then built by applying the weighted histogram analysis method (WHAM).⁽⁷¹⁾ Umbrella potentials were placed along the sampling coordinate varying from -180° to 180° (20 windows). The restraining potential was set to 35 kcal/mol/rad⁻¹ and the simulations were run in the NPT ensemble, using the same conditions as described above for the equilibrium trajectory.

3 | RESULTS AND DISCUSSION

For each of the anion considered, quantum chemistry calculations using two continuum models of solvation will be presented, followed by the calculations on clusters including explicit water molecules (Section 3.1). In Section 3.2, we shall describe the solvation properties that we investigated for the anions in bulk water by using a QM/MM scheme, providing a qualitative comparison with the picture stemming from static calculations.

3.1 | Results obtained using electronic structure calculations

| Succinate: SUC²⁻

We shall start the presentation of the results by providing a comparison with similar studies in the literature. In line with the work in Ref. 41, we found three conformers of succinate, the structure of which is reported in the top panel of Figure 2 along with some useful definitions. The conformers will be named SUC₁²⁻, SUC₂²⁻, SUC₃²⁻, corresponding to the S1, S2, S3 conformers obtained by Sutton et al. The results obtained by those authors were confirmed by performing a relaxed scan along the χ dihedral and reoptimizing the minima found along the constrained surface. A comparison of the minima sampled by this methodology using M05-2X/cc-pVDZ and SMD with the published structures.⁽⁴¹⁾ SUC₁²⁻ corresponds to a gauche conformation of the two carboxylates with respect to the bond formed by the two methylene carbons, whereas SUC₂²⁻ and SUC₃²⁻ correspond to an anti conformation and differ because of the relative orientation of the planes containing the COO⁻ groups.

The values of the C₁-C₂-C₃-C₄ dihedral as well as the relative energies with respect to SUC₁²⁻ are shown in Table 1. Three quantum chemistry levels were used as well as two different continuum models of solvation. The overall agreement among the different methods used in describing the most relevant features of the three conformers is satisfying. In the case of SUC₁²⁻, the value of the dihedral is slightly different in the case of PM6 calculations, and this methods describes a SUC₃²⁻ conformer that is practically isoenergetic to SUC₂²⁻. In addition, PM6+SMD predicts a

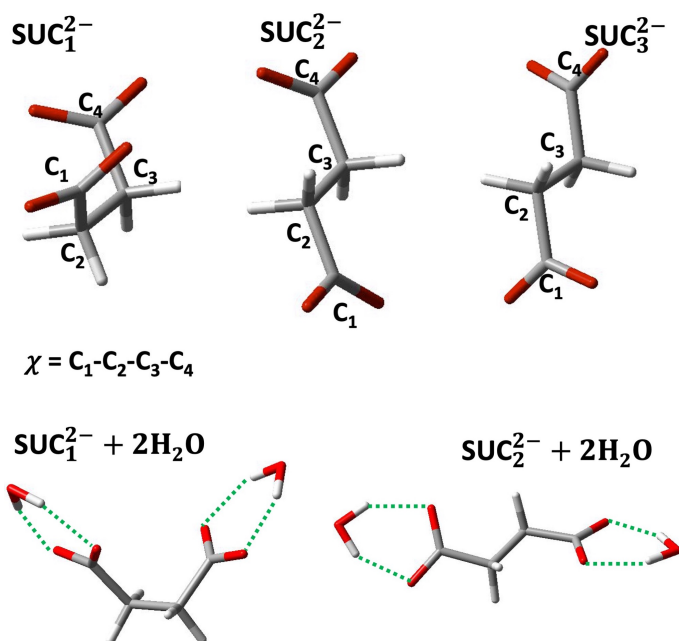


FIGURE 2 Top panel: structure of the three conformers of succinate and definition of the dihedral used in the analysis. SUC_1^{2-} , SUC_2^{2-} , SUC_3^{2-} correspond to the S1, S2, S3 conformers in Ref. 41, respectively. Bottom panel: structure of the optimized clusters formed by SUC_1^{2-} and SUC_2^{2-} and two water molecules next to the charged groups. Green dotted lines are shown as a guide for the eye along the succinate-water hydrogen bond directions.

169 higher energy for SUC_2^{2-} and SUC_3^{2-} than PM6+PCM. However, considering how close in energy such conformers are
 170 overall, the observed differences between two different quantum energy methods are of the same order of magnitude
 171 of the differences observed when the same QM method is used with two different solvation models. Notably, using
 172 M05-2X/cc-pVDZ+PCM predicts a lower energy SUC_3^{2-} conformers whereas the opposite is found when using M05-
 2X/cc-pVDZ+SMD.

TABLE 1 Summary of structural properties and relative energies for the SUC_1^{2-} , SUC_2^{2-} , SUC_3^{2-} conformers: the χ dihedral is defined in Figure 2. The relative energy is computed with respect to the energy of the SUC_1^{2-} conformer.

SUC_1^{2-}		
Method	χ	ΔE (kJ/mol)
B3LYP/6-311++G(d,p), PCM	-69°	0.0
B3LYP/6-311++G(d,p), SMD	-66°	0.0
M05-2X/cc-pVDZ, PCM	-69°	0.0
M05-2X/cc-pVDZ, SMD	-63°	0.0
PM6, PCM	-59°	0.0
PM6, SMD	-42°	0.0
SUC_2^{2-}		
Method	χ	ΔE (kJ/mol)
B3LYP/6-311++G(d,p), PCM	179°	-0.74
B3LYP/6-311++G(d,p), SMD	178°	+1.09
M05-2X/cc-pVDZ, PCM	177°	+1.06
M05-2X/cc-pVDZ, SMD	176°	+3.62
PM6, PCM	178°	+2.40
PM6, SMD	174°	+11.20
SUC_3^{2-}		
Method	χ	ΔE (kJ/mol)
B3LYP/6-311++G(d,p), PCM	-179°	-0.77
B3LYP/6-311++G(d,p), SMD	-176°	+1.30
M05-2X/cc-pVDZ, PCM	-162°	-0.51
M05-2X/cc-pVDZ, SMD	-173°	+4.19
PM6, PCM	-178°	+2.40
PM6, SMD	-173°	+11.20

173

174 We turn now to the discussion of the carboxyl stretch band in the infrared spectrum of succinate. A digitized
 175 Fourier Transform Infrared (FTIR) spectrum from Ref. 72 is shown in black in Figure 3, together with the results that
 176 we obtained from normal mode analysis based on quantum chemistry calculations. The Engauge Digitizer⁽⁷³⁾ was
 177 used to obtain the experimental spectrum. We compared unscaled results with those obtained using the following

178 scaling factors for the IR frequencies: 0.9495 for M05-2X/cc-pVDZ⁽⁷⁴⁾, 0.98 for PM6 (we adapted the scaling factor
179 proposed in Ref. 75) and 0.978 for B3LYP/6-311++G(d,p)⁽⁷⁶⁾. Intensities were rescaled to make them comparable
180 and were thus expressed in arbitrary units.

181 When looking at the different spectra for the anions in a continuum solvent, one basic observation is that the
182 computed spectra from conformers SUC_2^{2-} and SUC_3^{2-} display negligible differences. This is to be expected, since the
183 two structures are extremely similar. Two bands are observed in the computed spectra with one exception, namely
184 the PM6+SMD, for which the two bands merge. The analysis of normal modes allows to attribute the two bands to
185 symmetric (S, lower frequency) and antisymmetric (AS, higher frequency) stretch. In the comparison with experiment,
186 the B3LYP/6-311++G(d,p)+PCM spectra in water provide the best agreement, both for the position of the two peaks,
187 the frequency gap between them and the relative height of the peaks. Without using scaling factors, the two bands
188 are predicted with a red shift smaller than $\approx 5 \text{ cm}^{-1}$, whereas a slightly larger shift is observed when scaling is applied.
189 When using B3LYP/6-311++G(d,p)+SMD the gap between the peaks is too small, since the AS peak is strongly red
190 shifted, and this is particularly true for SUC_1^{2-} . Spectra obtained with M05-2X/cc-pVDZ+SMD were already discussed
191 in Ref. 41. They provide a reasonably good agreement with experiments when a scaling factor is applied but the band
192 gap is overestimated, a feature that is better reproduced with M05-2X/cc-pVDZ+PCM. Finally, the unscaled results
193 obtained with PM6+PCM are quite reasonable, though slightly blue shifted. In this case, scaling by using a 0.98 scaling
194 factor provides a much nicer agreement agreement with experiments.

195 The next step in our methodology was to include in our quantum chemistry calculations explicit water molecules
196 interacting directly with the charged groups. We then optimized one cluster containing two water molecules, one by
197 each carboxylate, for the three conformers considered. In all cases, starting from a cluster with SUC_3^{2-} led to a change
198 in conformation to SUC_2^{2-} , which is confirmed by the results in Ref. 41. The structures of the two optimized clusters
199 are reported in the bottom panel of Figure 2. The M05-2X/cc-pVDZ method predicts the SUC_2^{2-} cluster to be slightly
200 more stable than SUC_1^{2-} (by -1.10 kJ/mol with SMD and by -2.93 kJ/mol with PCM). The same trend is obtained
201 with PM6+SMD (SUC_2^{2-} more stable than SUC_1^{2-} by -5.66 kJ/mol), while the SUC_2^{2-} cluster is predicted to be more
202 stable by 0.68 kJ/mol using PM6+PCM, by 0.12 kJ/mol using B3LYP/6-311++G(d,p)+PCM and by 1.40 kJ/mol using
203 B3LYP/6-311++G(d,p)+SMD. As in the case of the isolated molecules in a continuum, such energy differences are
204 extremely small.

205 In both SUC_1^{2-} and SUC_2^{2-} minima, the water molecules coordinate to the carboxylate to form two hydrogen bonds
206 each, giving rise to structures that are visually very symmetric. For the two DFT methods with the SMD solvation
207 model, nor the value of the O-C-O angle ($125\text{-}126^\circ$) or that of the C-O distances ($1.26\text{-}1.27 \text{ \AA}$) are perturbed by the
208 formation of the succinate-water clusters. It is worth observing that, for each carboxylate, the two hydrogen bonds
209 do not form at the same distance: we measure 1.71 \AA and 2.27 \AA using M05-2X/cc-pVDZ+SMD and 1.73 \AA , 2.78
210 \AA using B3LYP/6-311++G(d,p)+SMD. The corresponding results are 1.70 \AA and 2.37 \AA for M05-2X/cc-pVDZ+PCM
211 and 1.67 \AA , 2.83 \AA for B3LYP/6-311++G(d,p)+PCM, respectively. Independently of the conformer, these hydrogen
212 bond distances are quite similar when using PM6+PCM (1.79 and 1.80 \AA), while they are different with PM6+SMD
213 (1.79 and 1.91 \AA). The O-C-O angle is about 120° and not much affected by the interaction with explicit water and
214 the two hydrogen bond distances are comparable ($1.79\text{-}1.80 \text{ \AA}$), with both PM6+PCM and PM6+SMD.

215 Moving to the infrared spectra for the microsolvation model, the S and AS peaks obtained for the two conformers
216 are reasonably similar in all cases. The result obtained with PM6+SMD is improved compared with the case in which
217 no explicit water molecule was taken into account with respect to the band position. The analysis of normal modes
218 shows that this is due to the water bend mode strongly coupling to the S stretch. On the other hand, in the case
219 of the M05-2X/cc-pVDZ+PCM results, the water bend absorbs between 1700 cm^{-1} and 1750 cm^{-1} , thus strongly
220 affecting the band shape for the AS stretch. Overall, the agreement with the experimental band is fairly good but

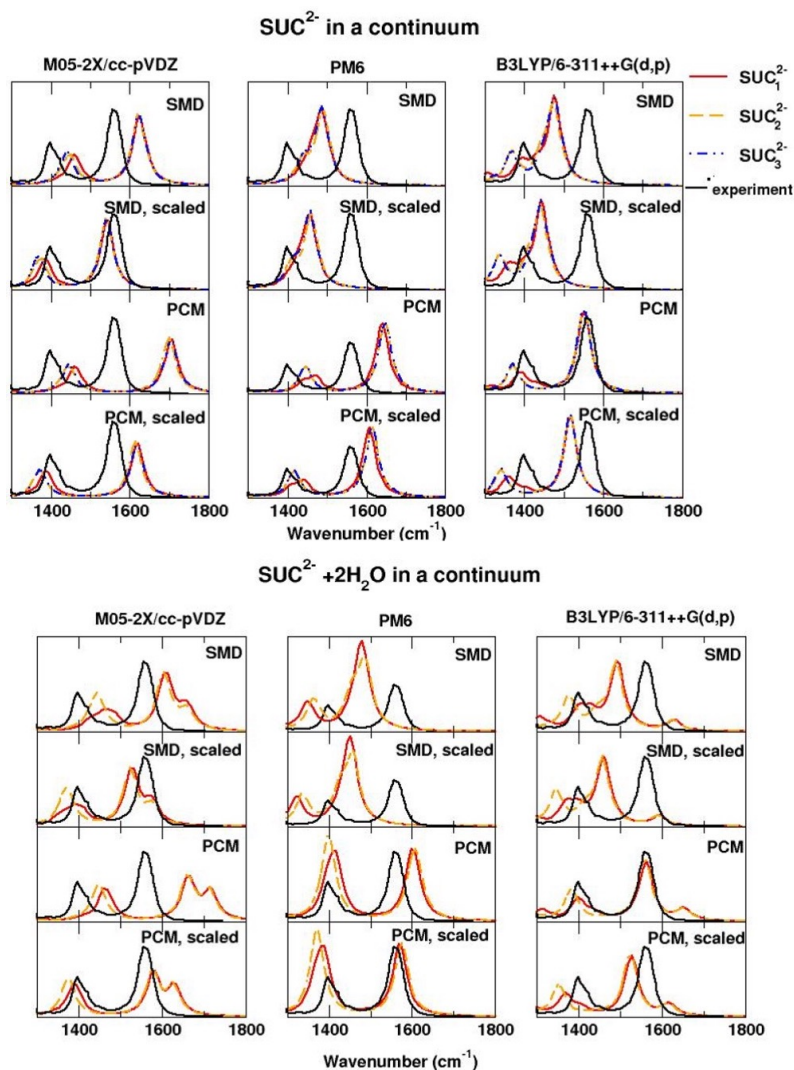


FIGURE 3 Infrared spectra obtained using different implicit solvation models (top panel) and incorporating two interacting water molecules (bottom panel). A comparison is provided with an experimental FTIR spectrum of succinate in water solution⁽⁷²⁾. Intensities are expressed in arbitrary units.

221 caution should be taken in discussing a direct comparison with the experiment, in which a difference spectrum was
 222 recorded. In principle, explicit contributions from the vibrational motions of water molecules (e.g. bending) is thus not
 223 present in experimental measurements.

224 | Aspartate: ASP^-

225 Considering the anions from the aspartic acid amino acid, in the case of ASP^- , recent work by Hernández et al.⁽⁴⁵⁾
 226 provided a deep analysis of the potential energy surface (PES) of the anion in a continuum solvent (water). By using
 227 the same DFT functional (B3LYP) and the same basis set (6-311++d(d,p)), we reproduced the same results using our
 228 approach, based on a relaxed scan along the $\text{C-C}_\alpha\text{-C}_\beta\text{-C}_\gamma$ dihedral and reoptimizing the minima thus found. The
 229 structure of the two conformers along with the atom numbering and the definition of some relevant geometrical
 230 coordinate are displayed in the top panel of Figure 4.

231 These two minima correspond roughly to an anti (ASP_1^-) and a gauche (ASP_2^-) conformation, considering the
 232 orientation of the carboxylate groups with respect to the $\text{C}_\alpha\text{-C}_\beta$ bond. The two conformers correspond to g-g+ and
 233 g-g- in Ref. 45, respectively. To describe the molecular structures, two more dihedrals were used in that work, named
 234 χ_1 and χ_2 . The values of the dihedral characterizing the conformers as well as the relative energies with respect to
 235 ASP_1^- are reported in Table 2. A comparison is provided between the PCM and SMD solvation models, and between
 236 the DFT level and the semiempirical PM6 results. Computations using M05-2X/cc-pVDZ led to an unphysical proton
 237 transfer from the charged amino group to one of the carboxylate oxygen atoms.

TABLE 2 Summary of structural properties and relative energies for the ASP_1^- and ASP_2^- conformers: the χ_i dihedrals are defined in Figure 4 and the relative energy is computed with respect to the energy of ASP_1^- .

ASP_2^-				
Method	χ	χ_1	χ_2	ΔE (kJ/mol)
B3LYP/6-311++G(d,p), PCM	-72°	49°	-34°,149°	+0.48
B3LYP/6-311++G(d,p), SMD	-65°	58°	-43°,140°	+2.26
PM6, PCM	-64°	56°	-43°,137°	-2.57
PM6, SMD	-51°	71°	-44°,136°	-7.05
ASP_1^-				
Method	χ	χ_1	χ_2	
B3LYP/6-311++G(d,p), PCM	-170°	-50°	-143°,38°	0.00
B3LYP/6-311++G(d,p), SMD	-179°	-57°	-146°,36°	0.00
PM6, PCM	-177°	-58°	-142°,39°	0.00
PM6, SMD	169°	-69°	-141°,40°	0.00

238 In both conformers, the stabilization of the optimized structures in water is linked with the formation of an in-
 239 tramolecular hydrogen bond between the acidic hydrogen atoms and the negatively charged carboxylate oxygen. DFT
 240 calculations predict ASP_2^- to be more stable than ASP_1^- while the opposite trend is predicted by PM6. As in the
 241 previous case, the differences in energies are very small. The structures predicted by the two methods are very similar
 242 with both solvation models, the only exception being the structure of ASP_2^- with SMD, for which the χ dihedral is

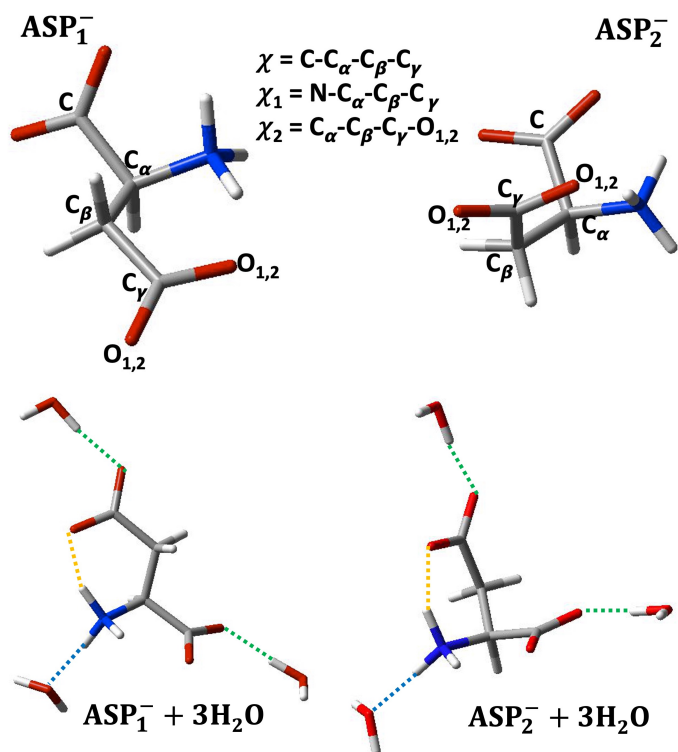


FIGURE 4 Top panel: structure of the two conformers of deprotonated aspartate and definition of the relevant dihedrals used in the analysis. Bottom panel: structure of the optimized clusters formed by ASP_1^- and ASP_2^- and three water molecules next to the charged groups. Relevant H-bond directions are displayed as a guide for the eye. Green dotted lines: intermolecular H-bonds, water is the donor. Blue dotted lines: intermolecular H-bond, water is the acceptor. Yellow dotted line: intramolecular H-bond.

243 different (about 10°).

244 In the top panel of Figure 5 we present a comparison of the computed carboxylate stretch bands using normal
245 mode analysis. A recorded spectrum of aspartic acid in water (ASP^- being the protonation state that is stable at neutral
246 pH) with a clear band assignment is not available in the literature and the difficulties in measuring such spectra in water
247 have been discussed in Ref. 77. Pearson and Slifkin⁽⁶⁵⁾ measured the IR difference spectra of amino acids in solutions
248 at different pH values but they mentioned that the peak positions and assignment might be affected by the strong
249 absorption of water in this region. At neutral pH, they assign a band at about 1400 cm^{-1} to S stretch and 1575 cm^{-1}
250 to asymmetric stretch. The S deformation of NH_3^+ occurs at about 1500 cm^{-1} , though it has been mentioned to occur
251 at 1474 cm^{-1} in another experimental work.⁽⁷⁸⁾ Using band deconvolution techniques, the authors of the latter study
252 assign two AS stretch vibrations, one for $\text{C}_\alpha\text{-C-O}$ (1601 cm^{-1}) and the other one for $\text{C}_\beta\text{-C-O}$ (1576 cm^{-1}). A recent
253 recorded spectrum for a 0.01 m solution of aspartic acid in water, appearing in Ref. 79, displays a strong band at 1600
254 cm^{-1} , a shoulder around 1510 cm^{-1} and another peak at 1350 cm^{-1} . However, band assignment is not provided
255 in the latter work. The PM6+SMD computed spectra display merged bands. All spectra tend to be blue shifted but
256 using scaling factors allows corrected spectra in a generally good agreement with experiments, the exception being
257 the spectrum predicted by using B3LYP/6-311++G(d,p)+SMD, which is red shifted. Normal modes analysis shows
258 that the NH_3^+ deformation couples strongly to the AS stretch and that, in addition, the lower energy tail of the band is
259 associated with a contribution from the C-H deformation. The position of the NH_3^+ does not seem to be well predicted.

261 Considering the general poorer agreement provided by SMD, to discuss the results obtained for the calculations
262 including explicit water molecules (structures displayed in the bottom panel of Figure 4), we shall focus on the results
263 obtained with PCM. For the ASP_2^- conformer and for B3LYP/6-311++G(d,p), there is no significant effect on the O-C-
264 O angles ($125\text{-}128^\circ$) upon formation of solute-water hydrogen bonds compared with the result obtained on the solute
265 in the continuum. In the absence of explicit water molecules, the two $\text{C}_\gamma\text{-O}$ bonds are different: the two distances
266 are equal to 1.24 and 1.26 Å. The longer distance is associated with the O atom involved in an intramolecular H-bond
267 with with one of the amino hydrogen atoms (bond displayed in yellow in the Figure, N...H distance: 1.61 Å).

268 Once the complex is formed, the N...H distance increases slightly (to 1.80 Å). Each of the carboxyl groups forms
269 one H-bond with one water, a situation that is different with respect to what was described for succinate. The C-O
270 distances are all equivalent and very little affected by the formation of H-bonds (1.25 Å). The O...H distances in the
271 presence of explicit water are equal to 1.73 Å. For the ASP_1^- conformer, the general trend is very similar to what was
272 discussed for ASP_2^- . The intramolecular hydrogen bond involving one amino H atoms is formed with one oxygen atom
273 bonded to the C carbon (instead of C_γ) and moves from 1.60 to 1.70 Å after interaction with explicit water molecules.
274 The results obtained with PM6 are very similar and will not be discussed in detail. The only relevant difference appears
275 for ASP_1^- , where both oxygen atoms of C_γ form H-bonds with one water molecule (H-bond distances: 1.90 and 1.88
276 Å). In addition, for both conformers, the $\text{H}_2\text{O}\dots\text{HN}$ distance is shorter (1.80 Å) than the one computed with DFT.

277 To conclude our discussion on the ASP^- conformers surrounded by explicit water molecules, we present com-
278 puted absorption spectra from normal mode analysis in the bottom panel of Figure 5. In the B3LYP/6-311G++(d,p)+PCM
279 spectrum, water bend is mostly coupling with AS stretch, whereas NH_3^+ deformation is coupling with the S stretch
280 band. As for the succinate spectra, the water bend vibration couples with AS stretches. In the PM6+PCM result,
281 the AS band is split and the normal mode analysis shows that this is due to the two carboxylate groups oscillating at
282 different frequencies. The NH_3^+ deformation couples with S stretches, giving rise to complex features on this band for
283 the B3LYP/6-311G++(d,p)+PCM result, but it appears as a distinct band in the PM6+PCM spectrum at about 1525
284 cm^{-1} (1494 cm^{-1} in the scaled spectrum). The C-H deformation at C_α and coupling with water band vibrations play
285 a role below 1400 cm^{-1} .

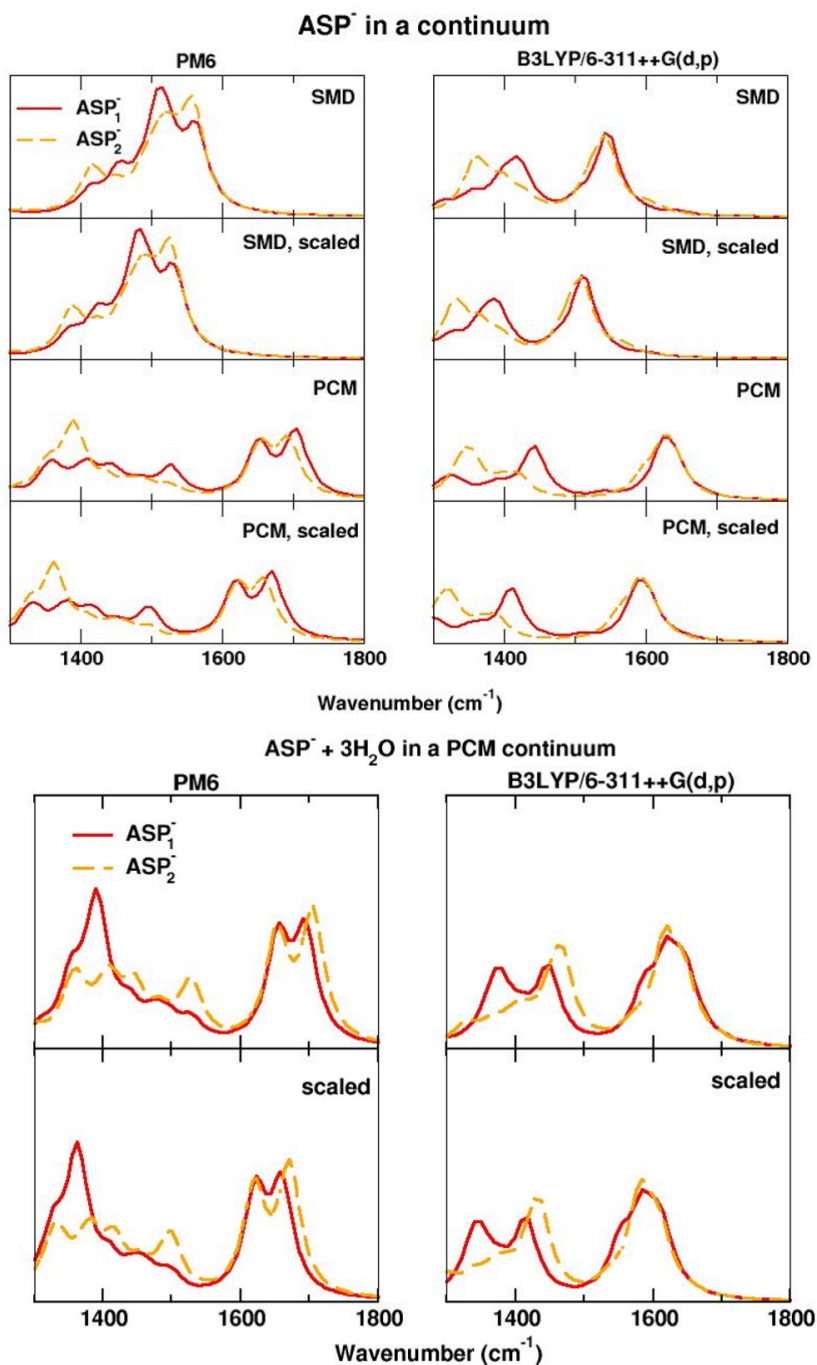


FIGURE 5 Infrared spectra for the two minima ASP₁⁻ and ASP₂⁻. Top panel: results for the anions in a continuum solvent. Bottom panel: results for the anions surrounded by three water molecules next to the polar groups in PCM. Intensities are expressed in arbitrary units.

286 | Aspartate: ASP^{2-}

287 Moving from ASP^- to ASP^{2-} , we find some similarities as well as some interesting differences. Two minima are found
 288 along the PES, named ASP_1^{2-} and ASP_2^{2-} , the structure of which is shown in the top panel of Figure 6. Following
 289 the same steps as for the previous anion, we provide results for DFT and PM6 calculations in SMD and in PCM
 290 water. Some of the relevant parameters to describe ASP_1^{2-} and ASP_2^{2-} are collected in Table 3. ASP_1^{2-} is predicted
 291 to be the most stable conformer with all methods but one exception, namely DFT+SMD. This might be related with
 292 a slightly different geometry predicted by this method for ASP_2^{2-} , displaying a somewhat different orientation of the
 293 carboxylates compared to the other methods. The same effect is observed with PM6+SMD and it also leads to a less
 294 stable ASP_1^{2-} compared to what is predicted by PM6+PCM. As for the other two anions, the energy differences are
 295 small. As a general remark, the carboxylate moieties are more free to reorient compared ASP_1^- and ASP_2^- because the
 296 formation an intramolecular H-bond is not as favorable as in those cases (the H atoms belonging to the amino site
 being less acidic).

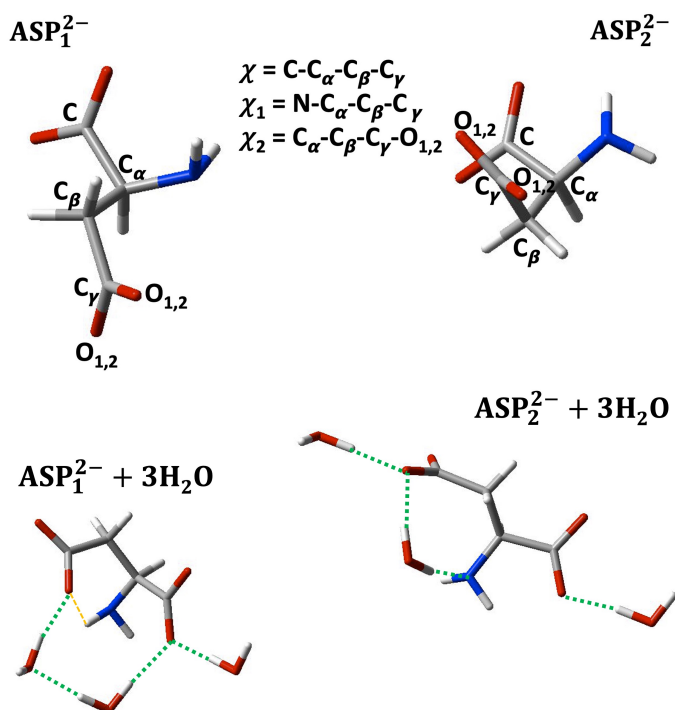


FIGURE 6 Top panel: structure of the two conformers of aspartate having a deprotonated amino group (ASP_1^{2-} , ASP_2^{2-}) and definition of the relevant dihedrals used in the analysis. Bottom panel: structure of the optimized clusters formed by ASP_1^{2-} and ASP_2^{2-} and three water molecules sitting next to the charged groups. Green dotted lines: intermolecular H-bonds, water is the donor. Yellow dotted line: intramolecular H-bond.

297

298 Wolpert and Helwig⁽⁷⁷⁾ have attributed the following frequencies, based on FTIR measurements of aqueous
 299 solutions of aspartic acid at pH = 10: the AS stretch band peaks at 1570 cm^{-1} and the S stretch at 1395 cm^{-1} . They
 300 also mention that each band presents a shoulder, a feature that they attribute to the different vibrations of each of the
 301 two carboxylates. Slightly different estimates were provided in Ref. 65: the S stretch band of the carboxylate group

TABLE 3 Summary of structural properties and relative energies for the ASP_1^{2-} and ASP_2^{2-} conformers: the χ_i dihedrals are defined in Figure 6 and the relative energy is computed with respect to the energy of the ASP_2^{2-} conformer.

ASP_1^{2-}				
Method	χ	χ_1	χ_2	ΔE (kJ/mol)
B3LYP/6-311++G(d,p), PCM	168°	-66°	-48°, 134°	+11.4
B3LYP/6-311++G(d,p), SMD	168°	-65°	-50°, 132°	-9.20
PM6, PCM	163°	-73°	-63°, 118°	+5.48
PM6, SMD	-179°	55°	-61°, 120°	+2.92
ASP_2^{2-}				
Method	χ	χ_1	χ_2	ΔE (kJ/mol)
B3LYP/6-311++G(d,p), PCM	-74°	53°	-126°, 55°	0.00
B3LYP/6-311++G(d,p), SMD	-65°	63°	-150°, 32°	0.00
PM6, PCM	-60°	65°	-63°, 118°	0.00
PM6, SMD	-62°	60°	-96°, 83°	0.00

302 was predicted at 1405 cm^{-1} , the AS stretch at 1575 cm^{-1} , the C-H bend and CH_2 wagging at about 1350 cm^{-1} , and no
 303 explicit band resulting from the N-H bend is expected in a basic environment in this frequency region. In qualitative
 304 agreement with experiments, in all the computed absorption spectra the contribution from the NH deformation is
 305 much smaller than in the ASP^- case. With PM6+SMD, the S and AS bands are very close (1450 and 1490 cm^{-1}) and
 306 the C-H deformation/wagging couple to both modes. With B3LYP/6-311++G(d,p)+SMD, the position of the bands
 307 is red shifted and the band gap is about 20 cm^{-1} too small. The scaled B3LYP/6-311++G(d,p)+PCM results for ASP_1^-
 308 provide the best agreement with the frequencies estimated from experiments (but the band gap is about 15 cm^{-1}
 309 larger, and it becomes too large when considering ASP_2^-). Coupling with the C-H deformation/wagging contributes to
 310 a large and intense AS stretch band. With PM6+SMD, the S and AS bands are very close, practically merging, and the
 311 C-H deformation/wagging couple to both modes. In the case of PM6+PCM, the band gap is larger and the mixing with
 312 N-H bend replaces the contribution from the methylene group. The scaled results provide a better agreement with
 313 experiments in this case. Considering the lack of a recorded spectrum and of a band assignment taking into account
 314 the band shape, and the small discrepancy between the available experimental results, it is not possible to push the
 315 comparison further. However, it is worth noting that mode dependent scaling factors have been proposed to achieve
 316 a quantitative agreement for the stretching band of the carboxylate groups of amino acids and peptides, though they
 317 were provided for different levels of theory.⁽⁸⁰⁾

318 In the clusters including three explicit water molecules (bottom panel of Figure 6), in the case of ASP_1^{2-} , a weak
 319 intramolecular H-bond is formed between one of the amino H atoms and one O atom on C_α (distance 2.35 \AA with
 320 DFT, 2.52 \AA with PM6). The N atom is not directly accepting a hydrogen bond from water, while one water molecule
 321 bridges between the other two interacting directly with the carboxylate O atoms. In the ASP_2^{2-} cluster, the intramolec-
 322 ular hydrogen bond is formed between HN and the O attached to C_α (distance: 2.01 \AA), whereas the N atom is ac-
 323 cepting a hydrogen bond from one water molecule (distance: 1.93 \AA with DFT, 2.09 \AA with PM6) that is also involved
 324 in a bond with a carboxylate (1.89 \AA - 1.83 \AA) in a bridging position. No apparent differences between the parameters
 325 describing the carboxylate groups are observed compared with the calculations on ASP_1^{2-} and ASP_2^{2-} in the absence

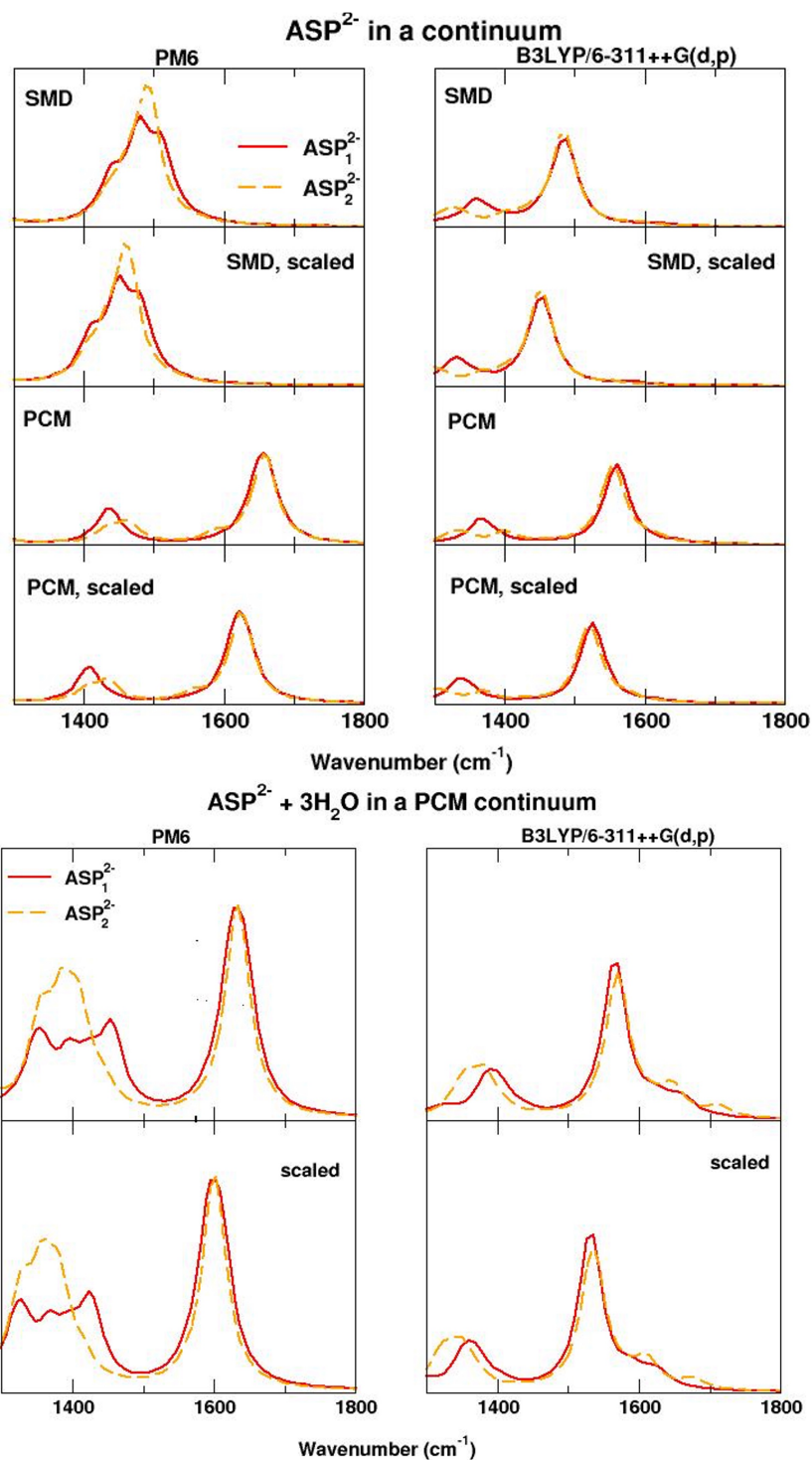


FIGURE 7 Infrared spectra for the two minima ASP_1^{2-} and ASP_2^{2-} . Top panel: results for the anions in a continuum solvent. Bottom panel: results for the anions surrounded by three water molecules next to the polar groups in PCM. Intensities are expressed in arbitrary units.

326 of explicit water.

327 Minor differences are observed in the IR spectra computed for the clusters to describe the carboxylate stretch
 328 band (bottom panel of Figure 7). Using B3LYP/6-311++G(d,p), we observe that the S stretch band is mostly affected
 329 by the the modes of the methylene group, whereas the AS stretch contains the coupling with the bend of the different
 330 water molecules. The higher frequency tail is due to the interaction with the N-H bend. Using PM6, the coupling of
 331 the water bend mode with the carboxylate stretch is possible for the S and the AS band, involving different water
 332 molecules. The contribution from the modes involving the methylene and amino groups are limited, compared with
 333 the results in a continuum.

334 3.2 | Results obtained using QM/MM molecular dynamics

335 We now move to a more realistic solvation model and consider a QM/MM simulation combining PM6 for the anions
 336 with TIP3P for the water molecules. We first focus on the conformational space and analyze the umbrella sampling
 337 profiles obtained by constraining the χ dihedral. Results are shown in Figure 8. For succinate, the global minimum
 338 is located at about 80° , corresponding to a distorted gauche conformations. Other, more shallow minima are present:
 339 one at -90° and the other corresponding to the anti conformation observed in SUC_2^{2-} and SUC_3^{2-} . The barriers to move
 340 from one basin to the other are quite small (of the order of 20-30 kJ/mol or smaller), we should therefore expect that
 341 the molecules explore different conformations along an MD trajectory. This is also true for aspartate, for both ASP^-
 342 and ASP^{2-} . However, the global minimum for ASP^- is located at -60° (corresponding to ASP_2^-), whereas the one for
 343 ASP^{2-} is located at about 55° (corresponding to ASP_2^{2-}). Given the limited sampling, we shall keep our analysis of the
 potential of mean force (PMF) to a qualitative description.

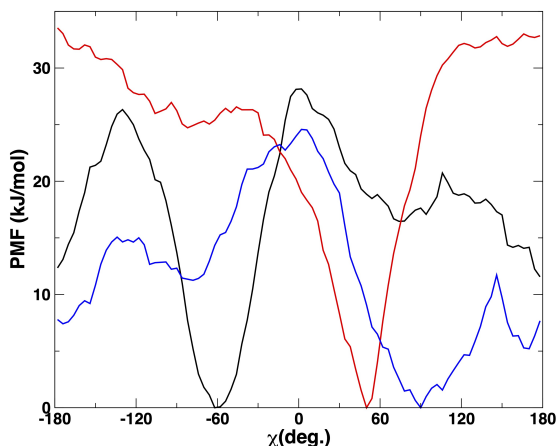


FIGURE 8 Potential of mean force obtained by umbrella sampling along the χ dihedral for the three molecules considered in this work: succinate (SUC_2^{2-} , blue curve), deprotonated aspartate (ASP^- , black curve), deprotonated aspartate in a basic environment (ASP_2^{2-} , red curve).

344

345 To show that our MD trajectory results are consistent with the PMF, we computed the distribution of the χ
 346 dihedrals measured along the dynamics. Results are reported as supporting information (Figure S1) and show that our
 347 analysis of the conformational space corresponds to basins explored along equilibrium MD.

348 The analysis of the solvation shell of the anions reveals interesting similarities with what has been obtained from
 349 electronic structure calculations. We shall start by describing the solute-solvent radial pair distribution functions
 350 (RDFs) that characterize the polar groups, displayed in Figure 9. For succinate, the interaction between all O atoms
 351 from the carboxylate groups with the H atoms of water is remarkably similar, which illustrate that our trajectory is
 352 long enough to sample the different solvation structures. The first peak, corresponding to the H-bond distance, is
 353 located at 1.83 Å. Two more, quite well defined, peaks are present, corresponding to the second and third hydration
 354 shell, hinting at a strong interaction, which is reasonable due to the charge on the carboxylate groups. Similar results
 355 are obtained for the two aspartate anions, for which a slightly less structured third peak is observed. However, in the
 356 case of ASP^- , we note a slight difference: one of the O...H RDFs curves displays a lower peak height for the first and
 357 second peaks. This behavior is related with one of the oxygen atoms of the carboxylate (belonging to C_γ) forming an
 358 intramolecular H-bond with one (two) of the H atoms amino group. One or two HN atoms being involved is related
 359 with the rotational motion of the NH_3^+ moiety along the dynamics. A distribution of intramolecular O...HN distances
 360 is reported as supporting information (Figure S2). This observation is in accord with what was found for ASP_2^- using
 361 electronic structure calculations.

362 When considering the interactions of the amino group with water, major differences are observed between the
 363 two aspartate anions. When the amino group is protonated (i.e. ASP^-), the H atoms are weakly donating a H-bond
 364 (the first peak is located at 2 Å), while the N atom does not interact directly with the H atoms of water. On the other
 365 hand, no interaction is observed between the HN atoms and water for the aspartate bearing a deprotonated amino
 366 group (ASP^{2-}), whereas the N atom behaves as an acceptor of H-bond. Again, this result is very well consistent with
 367 what was described for the quantum chemistry calculations including explicit water molecules.

368 The integration of the first peak of the RDFs up to the distance corresponding to the position of the first minimum
 369 r_{min} , according to the definition of coordination number $n_{r_{min}}$, can provide some insight on the number of water
 370 molecules surrounding the polar sites:

$$n_{r_{min}} = 4\pi\rho \int_0^{r_{min}} dr g(r) r^2 \quad (1)$$

371 For each O atom of the carboxylate groups in succinate we obtain 7 water hydrogen atoms. We also computed
 372 the distribution of angles formed between the direction of the water dipole moment and the C-O axis. We found
 373 an average angle of 60° for the maximum of the orientational correlation. This calculations, the results of which are
 374 reported as Supporting Information (Figure S3) is based on the normalized function⁽⁸²⁾

$$G_1(r) = \frac{h^{110}(r)}{3g_{000}(r)} = \langle \cos(\theta)_{12} \rangle, \quad (2)$$

375 where $h^{110}(r)$ is based on the orientational correlation function defined as

$$h^{110}(r) = \frac{\int h(12)\Phi^{110}(12)d\Omega_1\Omega_2}{\int [\Phi^{110}(12)]^2 d\Omega_1\Omega_2}, \quad (3)$$

376 with

$$\Phi^{110}(12) = \hat{u}_1 \cdot \hat{u}_2. \quad (4)$$

377 $h(12) = g(12) - 1$ is the total pair correlation function between molecules 1 and 2, Φ^{110} is the rotational invariant

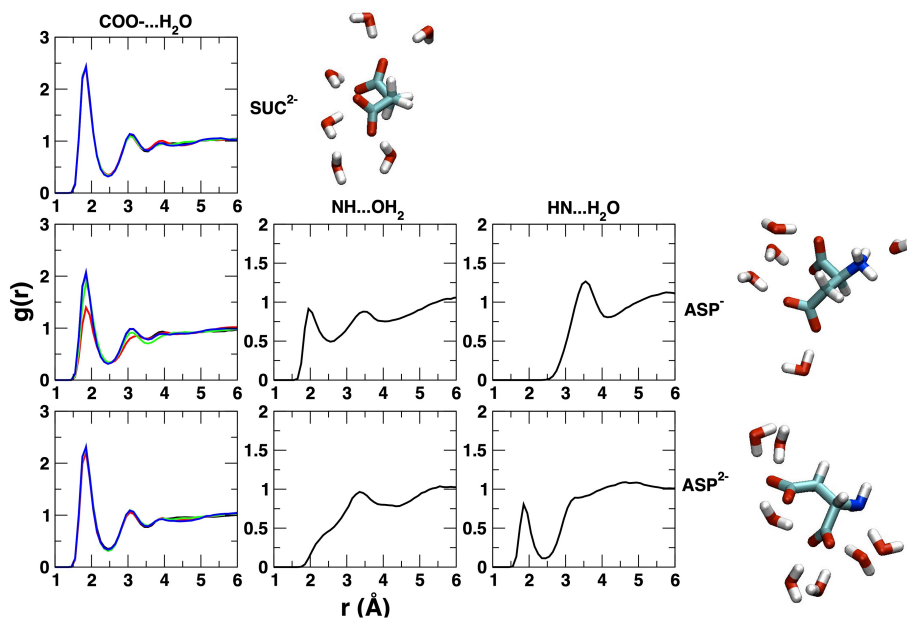


FIGURE 9 Relevant solute-water pair distribution functions. Top panel: interaction between oxygen atoms belonging to carboxylates in succinate (black, red, green, blue curves). Middle panel, from left to right: interaction between oxygen atoms belonging to carboxylates in ASP⁻ (black, red, green, blue curves); interaction with hydrogen atoms belonging to the amino group; interaction with the nitrogen atom belonging to the amino group. Bottom panel: same as for the middle panel, in the case of ASP²⁻. For each panels, images illustrating the structure of the solvation shell created by extracting snapshots along the simulated trajectories are displayed. The software used to create the images is VMD.⁽⁸¹⁾

378 function related with orientations, and Ω_i are the angular variables. We thus considered \hat{u}_1 as a unitary vector along
 379 one C-O axis and \hat{u}_2 as a unitary vector along the direction of the instantaneous molecular dipole of one solvent
 380 molecule.

381 This finding underlines that each O atom in one carboxyl group interacts with different water molecules, instead of
 382 interacting with the same water molecule, as in the optimized structures for the clusters involving SUC_1^{2-} and SUC_2^{2-}
 383 (for which the value of the angle is 45-50°). More importantly, in the latter cases one would expect a population of H-
 384 bonded water molecules at shorter distances (namely about 1.7 Å) and a sharper peak in the orientational correlations.
 385 As a matter of fact, this finding is no surprising, since the configuration adopted in bulk water leads to a smaller
 386 perturbation of the local H-bond network in water compared with the one found for the clusters, and is compatible
 387 with the action of molecular motions that are active when including temperature effects.

388 The orientation of water molecules with respect to the C-O axes of the carboxylate groups is only slightly different
 389 between succinate and the two aspartate anions considered, since we obtained a value of 66° associated with the
 390 first peak of the G_1 function. From the RDFs, the number of water H atoms surrounding these groups is about 7, as
 391 in the case of succinate, with the exception of the C-O bond that is involved in an intramolecular hydrogen bond (in
 392 ASP^-), for which only two H atoms were found. The H-bond donor NH_3^+ moiety in ASP^- is surrounded by 5 H atoms.

393 The infrared absorption spectrum of a molecule can be evaluated based on MD trajectories by using a formulation
 394 stemming from Kubo's relationship⁽⁸³⁾. According to this development, the infrared intensity can be expressed in
 395 terms of the Fourier transform of the molecular dipole time correlation function as:

$$I(\omega) = \frac{1}{2\pi} Q(\omega) \int_{-\infty}^{+\infty} dt \exp(-i\omega t) \langle \mu(0) \cdot \mu(t) \rangle \quad (5)$$

396 in which $Q(\omega)$ is a prefactor taking into account the quantum nature of the time correlation function.^(84,85) An explicit
 397 treatment of this factor is not included in our calculation since this goes beyond the purposes of our analysis. We used
 398 two methods to compute numerically the infrared spectrum by using the instantaneous molecular dipole provided by
 399 the QM output in the QM/MM setup: a method based on the Wiener-Khinchin theorem and the maximum entropy
 400 method, implemented by using the routines provided in Ref. 86. The power spectra thus obtained are displayed in
 401 Figure 10.

402 The carboxylate AS stretch band peaks at 1670 cm^{-1} , compared with 1560 cm^{-1} in experiments. The intensity
 403 of the S band is quite low and the band seems very broad, compared with experiments⁽⁷²⁾ (for which the peak is
 404 located at 1400 cm^{-1}). From our electronic structure PM6 calculations, we may notice that the AS stretch band is
 405 more sensitive to the solvation model used, whereas the S stretch band is located at about 1400 cm^{-1} in all cases.
 406 Peaks at lower frequencies than 1400 cm^{-1} are due to C-H bend motions, so that we can safely assume that the peak
 407 on the lower frequency side of the top panel in Figure 10 is due to this mode.

408 In the case of the computed band for ASP^- , the AS stretch band is larger but still peaking at 1670 cm^{-1} . As it was
 409 already discussed, the experiments suggest that the AS stretch can couple with the bend modes of the NH_3^+ group,
 410 and also in this case its position is strongly dependent on the solvation model used. The lower frequency portion of
 411 the S stretch band that is observed in the spectrum is most likely due to the CH_2 modes. The results are very similar
 412 in the case of ASP^{2-} but the AS band appears narrower, which is consistent with the experimental interpretation
 413 based on a lower coupling with the NH bend.⁽⁶⁵⁾ To conclude with, it is worth noting that, in the case of aspartate,
 414 the QM/MM MD based results display band shapes and positions that are very similar to the results obtained on the
 415 anions with PM6+PCM.

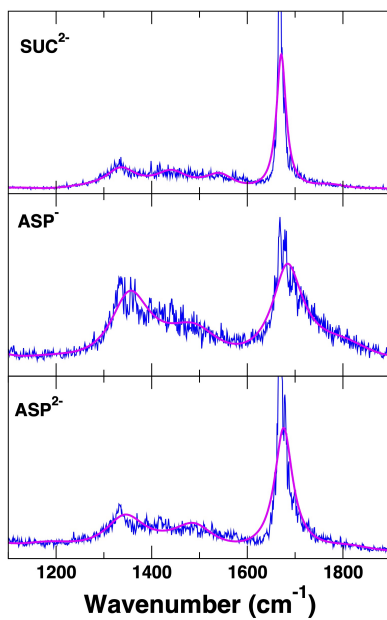


FIGURE 10 Infrared absorption spectra of the three anions considered. Magenta line: result based on the maximum entropy method. Blue line: result based on the Wiener-Khinchin theorem. Intensities are expressed in arbitrary units.

4 | CONCLUSION

In this work, we provided a bottom-up computational approach to study the conformational space and the vibrational properties of three anions containing carboxylate groups: succinate, zwitterionic aspartate and aspartate having a deprotonated amino group. More specifically, we focused on the carboxyl stretch band, which is sensitive to the environment and therefore can be used as a probe of local structure and of solvation, and on an exploration of the conformational space of the molecules interacting with water. In the case of succinate and zwitterionic aspartate, some previous computational work in the literature provided important results to compare with.^(41,45) Similarly to what had been done in those studies, we started from the anions in a continuum solvent, then we moved to a (simplified) microsolvation approach, in which we included one water molecule sitting next to each polar group and immersed the cluster thus obtained in a continuum solvent. In addition, we performed hybrid QM/MM MD simulations of the anions in bulk water, thus including finite temperature effects.

In the electronic structure calculations, two DFT methods (M05-2X/ccpVDZ and B3LYP/6-311++G(d,p)) and one semiempirical method (PM6) were employed, and for each method we assessed the performance of two solvation models (SMD, PCM). The results on the predicted geometries, both for the anions and the clusters, were remarkably similar. This might result surprising in the case of the clusters, considering the artifacts stemming from applying semiempirical Hamiltonians to describe non-covalent interactions.⁽⁸⁷⁾ However, we can hypothesize that the strong electrostatic interactions between negative charged anions (possessing polar groups) and water can be more easily modeled than the interaction between neutral systems, since the limitations of semiempirical methods are more apparent when weak non-covalent interactions are present.

When comparing the results obtained for the carboxyl stretch band based on normal mode analysis, we found that the overall agreement with experimental measurements is promising. However, for the considered anions the PCM outperforms SMD in the prediction of both the position and the AS/S stretch band gap. Using PCM, in the comparison between scaled and unscaled frequencies, we showed that the results obtained with PM6 are systematically improved, leading to an excellent agreement with the available experimental results. Remarkably good results are obtained with B3LYP/6-311++G(d,p)+PCM but the use of a scaling factor in this case does not necessarily lead to an improvement, since non-scaled frequencies can be within 5 cm^{-1} with respect to experiments, as in the case of succinate (the only anion for which a recorded spectrum in water was available in the literature). We can thus conclude that the PCM solvation model with B3LYP/6-311++G(d,p) can provide quantitative agreement with experimental measurements of the carboxylate stretch band and that, notwithstanding the lower level of quantum chemistry, scaled PM6+PCM spectra can lead to a very close performance. One additional, interesting conclusion that we can draw is that one observes larger differences when comparing results obtained with the same quantum chemistry method but a different continuum solvation model, than when comparing two different quantum chemistry methods. Such finding should warn about carrying out calculations of such properties without careful benchmarking. In addition, it has been pointed out⁽⁸⁰⁾ that scaling factors might depend on the vibrational motion in floppy biomolecules, which would need some additional work.

A description of the local structuring of water molecules around the anions was achieved though the analysis of simulated trajectories in the PM6/MM hybrid scheme. In all treated systems, each carboxylate group interacts with different water molecules, a configuration that leads to a limited perturbation of the H-bond network in water. The H atoms belonging to the charged amino group in the zwitterionic aspartate act as H-bond donors and an intramolecular H-bond, already characterized through QM calculations, is also present. On the other hand, the amino group acts as an H-bond acceptor when it is deprotonated. An exploration of the conformational space of the anions by umbrella sampling along the C-C-C-C dihedral yielded a very similar description compared with the results from electronic

458 structure calculation and it was consistent with the distribution of the dihedral along an equilibrium trajectory. The
459 IR band simulated by computing the power spectrum of the QM dipole delivers the same picture as the IR band from
460 QM calculations. Further improvement on the band shape and (sub)band position would require a specific treatment
461 of quantum effects^(84,85) and/or more advanced techniques,⁽⁴⁴⁾ which is beyond the scope of the present work.

462 In future and ongoing work, we are interested in moving to anions containing carboxylates intercalated into
463 LDH. The questions that we would like to answer involve a molecular understanding of the balance between 1) anion-
464 surface, 2) anion-water, 3) water-surface interactions and 4) of nanoconfinement effects. The present study has
465 brought a detailed knowledge on how to treat the effect of H-bonds formed by carboxylate units, which can be
466 exploited in the analysis of LDH, since anchoring of the anion on the inorganic surface is mediated by the formation
467 of hydrogen bonds with the hydroxyl groups. Competition between different systems capable of forming H-bonds
468 (surface, anions, water) is therefore expected to play an important role and is likely to be sensitive to the hydration
469 content.

470 ACKNOWLEDGMENTS

471 The authors would like to thank the local HPC cluster at LPCT for computing time and Dr. Fabien Pascale for skillful
472 assistance. V. K. P. gratefully acknowledges the University of Lorraine and the C2MP PhD school for a PhD grant.

473 Conflict of interest

474 The authors declare no conflict of interest.

475 SUPPORTING INFORMATION

476 Supporting Information is available. Additional data from this study are available from the corresponding author upon
477 request.

478 REFERENCES

- 479 [1] B.-H. Ye, M.-L. Tong, X.-M. Chen, *Coord. Chem. Rev.* **2005**, *249* (5-6), 545–565.
- 480 [2] S. L. James, *Chem. Soc. Rev.* **2003**, *32* (5), 276–288.
- 481 [3] X.-X. Xie, Y.-C. Yang, B.-H. Dou, Z.-F. Li, G. Li, *Coord. Chem. Rev.* **2020**, *403*, 213100.
- 482 [4] C. Rao, S. Natarajan, R. Vaidhyanathan, *Angew. Chem. Int. Ed.* **2004**, *43* (12), 1466–1496.
- 483 [5] F. S. Richardson, *Chem. Rev.* **1982**, *82* (5), 541–552.
- 484 [6] A. Thibon, V. C. Pierre, *Anal. Bioanal. Chem.* **2009**, *394* (1), 107–120.
- 485 [7] L. Ackermann, *Acc. Chem. Res.* **2014**, *47* (2), 281–295.
- 486 [8] M. Grätzel, *J. Photochem. Photobiol. C Rev.* **2003**, *4* (2), 145–153.
- 487 [9] B. Das, A. Rahaman, A. Shatskiy, O. Verho, M. D. Karkas, B. Åkermark, *Acc. Chem. Res.* **2021**, *54* (17), 3326–3337.
- 488 [10] T. Dudev, C. Lim, *J. Phys. Chem. B* **2004**, *108* (14), 4546–4557.

- 489 [11] T. Dudev, C. Lim, *Acc. Chem. Res.* **2007**, 40 (1), 85–93.
- 490 [12] H. Dong, J. F. Snyder, K. S. Williams, J. W. Andzelm, *Biomacromolecules* **2013**, 14 (9), 3338–3345.
- 491 [13] A. M. Du Poset, A. Zitolo, F. Cousin, A. Assifaoui, A. Lerbret, *Phys Chem Chem Phys* **2020**, 22 (5), 2963–2977.
- 492 [14] A. Assifaoui, A. Lerbret, H. T. Uyen, F. Neiers, O. Chamin, C. Loupiac, F. Cousin, *Soft Matter* **2015**, 11 (3), 551–560.
- 493 [15] I. Imaz, M. Rubio-Martinez, J. An, I. Sole-Font, N. L. Rosi, D. Maspocho, *Chem. Comm.* **2011**, 47 (26), 7287–7302.
- 494 [16] C. Geffroy, A. Foissy, J. Persello, B. Cabane, *J. Colloid Interface Sci.* **1999**, 211 (1), 45–53.
- 495 [17] Y. Lu, J. D. Miller, *J. Colloid Interface Sci.* **2002**, 256 (1), 41–52.
- 496 [18] D. C. Standnes, T. Austad, *J. Pet. Sci. Eng.* **2000**, 28 (3), 123–143.
- 497 [19] R. Zhang, N. Qin, L. Peng, K. Tang, Z. Ye, *Appl. Surf. Sci.* **2012**, 258 (20), 7943–7949.
- 498 [20] J. Kubelka, S. Bai, M. Piri, *J. Phys. Chem. B* **2021**, 125 (4), 1293–1305.
- 499 [21] C. Forano, T. Hibino, F. Leroux, C. Taviot-Guého, *Dev. Clay Sci.* **2006**, 1, 1021–1095.
- 500 [22] D. G. Evans, R. C. Slade, *Layered double hydroxides* **2006**, 1–87.
- 501 [23] L. Mohapatra, K. Parida, *J. Mater. Chem. A* **2016**, 4 (28), 10744–10766.
- 502 [24] C. Del Hoyo, *Appl. Clay Sci.* **2007**, 36 (1-3), 103–121.
- 503 [25] G. Mishra, B. Dash, S. Pandey, *Appl. Clay Sci.* **2018**, 153, 172–186.
- 504 [26] K.-H. Goh, T.-T. Lim, Z. Dong, *Water Res.* **2008**, 42 (6-7), 1343–1368.
- 505 [27] S. Carlino, *Solid State Ion* **1997**, 98 (1-2), 73–84.
- 506 [28] J. Zhang, F. Zhang, L. Ren, D. G. Evans, X. Duan, *Mater. Chem. Phys.* **2004**, 85 (1), 207–214.
- 507 [29] N. T. Whilton, P. J. Vickers, S. Mann, *J. Mater. Chemistry* **1997**, 7 (8), 1623–1629.
- 508 [30] H. Nakayama, N. Wada, M. Tshukako, *Int. J. Pharm.* **2004**, 269 (2), 469–478.
- 509 [31] S. Aisawa, S. Takahashi, W. Ogasawara, Y. Umetsu, E. Narita, *J. Solid State Chem.* **2001**, 162 (1), 52–62.
- 510 [32] J. D. Bernal, *Proc. Phys. Soc. B* **1949**, 62 (10), 597.
- 511 [33] A. Meister, *Biochemistry of the amino acids*, Elsevier, **2012**.
- 512 [34] V. Erastova, M. T. Degiacomi, D. G. Fraser, H. C. Greenwell, *Nature Comm.* **2017**, 8 (1), 1–9.
- 513 [35] P. P. Kumar, A. G. Kalinichev, R. J. Kirkpatrick, *J. Phys. Chem. C* **2007**, 111 (36), 13517–13523.
- 514 [36] A. G. Kalinichev, P. Padma Kumar, R. James Kirkpatrick, *Philos. Mag.* **2010**, 90 (17-18), 2475–2488.
- 515 [37] B. Grégoire, V. Erastova, D. L. Geatches, S. J. Clark, H. C. Greenwell, D. G. Fraser, *Geochim. Cosmochim. Acta* **2016**, 176,
516 239–258.
- 517 [38] A. Tsukanov, S. Psakhie, *Sci. Rep.* **2016**, 6 (1), 1–8.
- 518 [39] G. Deacon, R. Phillips, *Coord. Chem. Rev.* **1980**, 33 (3), 227–250.

- 519 [40] V. Otero, D. Sanches, C. Montagner, M. Vilarigues, L. Carlyle, J. A. Lopes, M. J. Melo, *J. Raman Spectrosc.* **2014**, *45* (11-12),
520 1197–1206.
- 521 [41] C. C. Sutton, G. V. Franks, G. da Silva, *Spectrochim Acta A Mol Biomol Spectrosc* **2015**, *134*, 535–542.
- 522 [42] M. Nara, H. Torii, M. Tasumi, *J. Phys. Chem.* **1996**, *100* (51), 19812–19817.
- 523 [43] C. C. Sutton, G. da Silva, G. V. Franks, *Chem. Eur. J.* **2015**, *21* (18), 6801–6805.
- 524 [44] S. M. V. Pinto, N. Tasinato, V. Barone, A. Amadei, L. Zanetti-Polzi, I. Daidone, *Phys. Chem. Chem. Phys.* **2020**, *22* (5),
525 3008–3016.
- 526 [45] B. Hernández, F. Pflüger, M. Ghomi, *J. Comp. Chem.* **2020**, *41* (14), 1402–1410.
- 527 [46] J. J. Stewart, *J. Mol. Mod.* **2007**, *13* (12), 1173–1213.
- 528 [47] G. M. Torrie, J. P. Valleau, *J. Comput. Phys.* **1977**, *23* (2), 187–199.
- 529 [48] J. H. Jensen, M. S. Gordon, *Journal of the American Chemical Society* **1995**, *117* (31), 8159–8170.
- 530 [49] D.-S. Ahn, S.-W. Park, I.-S. Jeon, M.-K. Lee, N.-H. Kim, Y.-H. Han, S. Lee, *The Journal of Physical Chemistry B* **2003**, *107*
531 (50), 14109–14118.
- 532 [50] C. O. da Silva, B. Mennucci, T. Vreven, *The Journal of Physical Chemistry A* **2003**, *107* (34), 6630–6637.
- 533 [51] H.-S. Kim, D.-S. Ahn, S.-Y. Chung, S. K. Kim, S. Lee, *The Journal of Physical Chemistry A* **2007**, *111* (32), 8007–8012.
- 534 [52] S. M. Bachrach, *The Journal of Physical Chemistry A* **2008**, *112* (16), 3722–3730.
- 535 [53] E. Tang, D. Di Tommaso, N. H. de Leeuw, *Physical Chemistry Chemical Physics* **2010**, *12* (41), 13804–13815.
- 536 [54] D. Scuderi, J. Bakker, S. Durand, P. Maitre, A. Sharma, J. Martens, E. Nicol, C. Clavaguera, G. Ohanessian, *International*
537 *Journal of Mass Spectrometry* **2011**, *308* (2-3), 338–347.
- 538 [55] C. García-Iriepa, M. Zemmouche, M. Ponce-Vargas, I. Navizet, *Physical Chemistry Chemical Physics* **2019**, *21* (8), 4613–
539 4623.
- 540 [56] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A.
541 Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P.
542 Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng,
543 A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara,
544 K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A.
545 Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A.
546 Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M.
547 Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox,
548 *Gaussian~16 Revision C.01*, **2016**, Gaussian Inc. Wallingford CT.
- 549 [57] R. Dennington, T. A. Keith, J. M. Millam, *GaussView Version 6*, **2019**, Semichem Inc. Shawnee Mission KS.
- 550 [58] Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theory Comput.* **2006**, *2* (2), 364–382.
- 551 [59] T. H. Dunning Jr, *J. Chem. Phys.* **1989**, *90* (2), 1007–1023.
- 552 [60] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98* (45), 11623–11627.
- 553 [61] A. McLean, G. Chandler, *J. Chem. Phys.* **1980**, *72* (10), 5639–5648.
- 554 [62] R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople, *J. Chem. Phys.* **1980**, *72* (1), 650–654.

- 555 [63] C. J. Cramer, D. G. Truhlar, *Acc. Chem. Res.* **2008**, *41* (6), 760–768.
- 556 [64] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105* (8), 2999–3094.
- 557 [65] J. Pearson, M. Slifkin, *Spectrochim Acta A Mol Biomol Spectrosc* **1972**, *28* (12), 2403–2417.
- 558 [66] D. Case, R. Betz, D. Cerutti, T. Cheatham, III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi,
559 P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz,
560 G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-
561 Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, P. Kollman, *AMBER 2016*, **2016**, University of California, San
562 Francisco.
- 563 [67] H. C. Andersen, *J. Chem. Phys.* **1980**, *72* (4), 2384–2393.
- 564 [68] H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.
- 565 [69] K. Nam, J. Gao, D. M. York, *J. Chem. Theory Comput.* **2005**, *1* (1), 2–13.
- 566 [70] D. J. Price, C. L. Brooks III, *J. Chem. Phys.* **2004**, *121* (20), 10096–10103.
- 567 [71] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, *J. Comp. Chem.* **1995**, *16* (11), 1339–1350.
- 568 [72] N. J. Nicholas, G. V. Franks, W. A. Ducker, *Langmuir* **2012**, *28* (18), 7189–7196.
- 569 [73] *Engauge Digitizer*, <https://markumitchell.github.io/engauge-digitizer/>.
- 570 [74] M. L. Laury, S. E. Boesch, I. Haken, P. Sinha, R. A. Wheeler, A. K. Wilson, *J. Comp. Chem.* **2011**, *32* (11), 2339–2347.
- 571 [75] Z. A. Fekete, E. A. Hoffmann, T. Körtvélyesi, B. Penke, *Mol. Phys.* **2007**, *105* (19–22), 2597–2605.
- 572 [76] S. Jarmelo, D. Marques, P. Simões, R. Carvalho, C. Batista, C. Araujo-Andrade, M. Gil, R. Fausto, *J. Phys. Chem. B* **2012**,
573 *116* (1), 9–21.
- 574 [77] M. Wolpert, P. Hellwig, *Spectrochim Acta A Mol Biomol Spectrosc* **2006**, *64* (4), 987–1001.
- 575 [78] S. Y. Venyaminov, N. Kalnin, *Biopolymers* **1990**, *30* (13–14), 1243–1257.
- 576 [79] K. Dhal, S. Singh, M. Talukdar, *J. Mol. Liquids* **2022**, *352*, 118659.
- 577 [80] Y. Bouteiller, J.-C. Pouilly, G. Grégoire, *Comput. Theor. Chem.* **2011**, *966* (1–3), 220–224.
- 578 [81] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38.
- 579 [82] G. Stell, G. Patey, J. Høye, *Adv. Chem. Phys., Volume 48* **1981**, 183–328.
- 580 [83] D. A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, CA, **2000**.
- 581 [84] S. Egorov, J. Skinner, *Chem. Phys. Lett.* **1998**, *293* (5–6), 469–476.
- 582 [85] R. Ramirez, T. López-Ciudad, P. Kumar P, D. Marx, *J. Chem. Phys.* **2004**, *121* (9), 3973–3983.
- 583 [86] W. H. Press, S. A. Teukolsky, W. R. Vetterling, B. P. Flannery, *Numerical Recipes in Fortran*, Cambridge University Press;
584 Cambridge, UK, **1992**.
- 585 [87] A. Marion, G. Monard, M. F. Ruiz-López, F. Ingrosso, *J. Chem. Phys.* **2014**, *141* (3), 07B615_1.

586 GRAPHICAL ABSTRACT

587

588

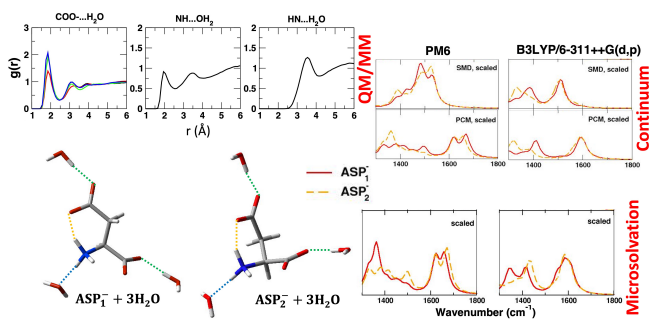
589

590

591

592

593



Structural and vibrational properties of carboxylates are modeled by using a bottom-up approach. Larger differences are found using different continuum solvation models and the same level of quantum chemistry compared with what is found using two different levels and the same solvation scheme.

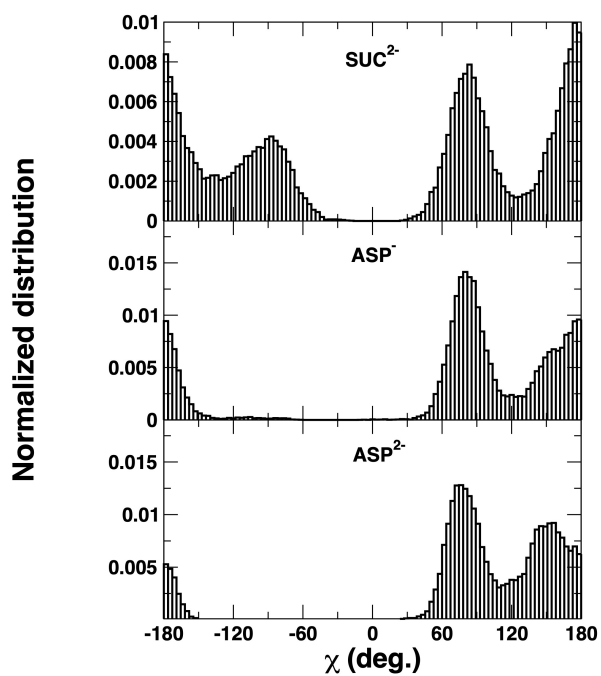


FIGURE S1 Normalized distribution of two relevant dihedral characterizing the conformations of the three anions considered in water.

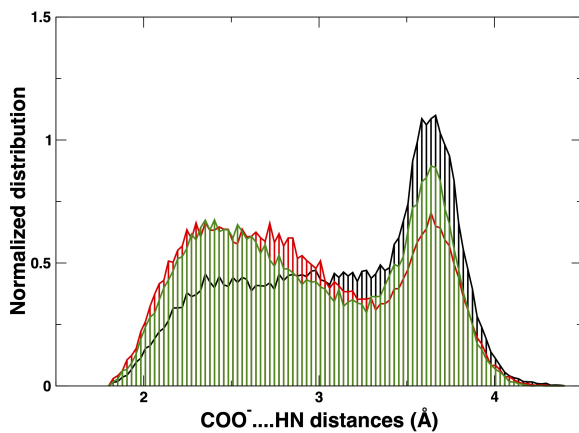


FIGURE S2 Normalized distribution of the $C_{\gamma} \text{OO}^{-} \dots \text{HNH}_2^{+}$ distances corresponding to the formation of an intramolecular hydrogen bond in aspartate (ASP^{-}).

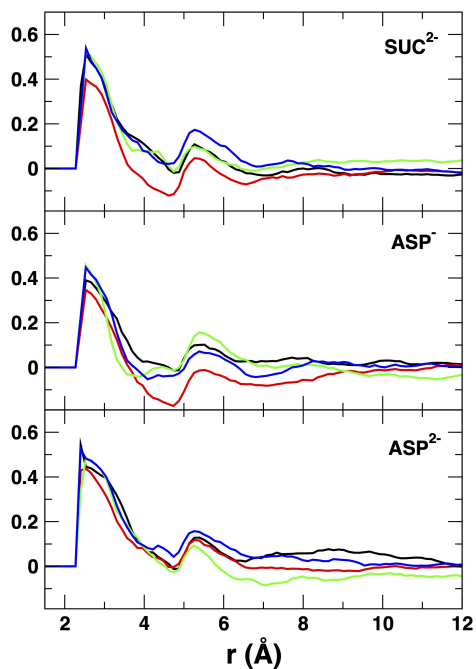


FIGURE S3 Orientational correlations based on the calculation of the $G_1(r)$ normalized function. The different colors refer to the four different O atoms of the carboxylate groups.