Journal of Molecular Liquids 319 (2020) 114109

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Free energy gradient for understanding the stability and properties of neutral and charged L-alanine molecule in water

Valdemir Ludwig [a,*], Zélia M. da Costa Ludwig [a], Danillo Valverde [b], Herbert C. Georg [c], Sylvio Canuto [b]

[a] Departamento de Física, Universidade Federal de Juiz de Fora, CEP 36036-330 Juiz de Fora, MG, Brazil
[b] Instituto de Física, Universidade de São Paulo, Rua do Matão 1371, Cidade Universitária, CEP 05508-090, SP, Brazil
[c] Instituto de Física, Universidade Federal de Goiás, CEP 74001-970 Goiânia, GO, Brazil

ARTICLE INFO

Article history:
Received 21 May 2020
Received in revised form 5 August 2020
Accepted 19 August 2020
Available online 25 August 2020

ABSTRACT

Amino acids exhibit their most important properties in aqueous environment in physiological conditions. These properties are related to their structural changes in water. Although several studies have been performed earlier for obtaining the structure of alanine in a solvent environment, an explicit consideration of the solvent molecular system is still in need. This work uses a Free Energy Gradient method to obtain the structures of alanine with an explicit consideration of the water environment. This is combined with a QM/MM method. Next, the electronic and vibrational properties of L-Alanine in water solution at different pH scales are obtained. The equilibrium structures of L-alanine are obtained in its neutral, anionic and cationic form. A systematic study of the solute-solvent hydrogen bonds was performed to predict the stability of neutral and charged L-alanine molecules. Quantum mechanical calculations were also performed to obtain the infrared spectrum, pH induced changes and to predict NMR parameters such as chemical shifts, magnetic shieldings and spin-spin coupling constants.

1. Introduction

In physiological conditions, most of biomolecules like amino acids, peptides and proteins exist predominantly in the zwitterionic (ZW) form. However, the zwitterionic amino acids are not stable in the gas phase and this excludes the possibility of understanding their biological functions [1,2] considering the isolated molecule. The stability and the interaction with water molecules is the first step for understanding of their structures, properties, and biological functions. On the other hand, the structure of amino acids is sensitive to the pH regulation which further affect the functionalities of living systems. Additionally, the pH of the liquid solution leads to large differences in the stability and electronic properties of amino acids, which implies that positively charged forms are prevalent in acidic environments and anionic forms prevail in basic solutions. Experimental NMR measurements [3–5] of L-Alanine and L-Alanyl-L-Alanine at various pH aqueous environments measured a strong interaction between aqueous solution and molecular flexibility of the solute. Furthermore, a detailed band analyses was made in the IR spectra of experimental samples of different protonation states of solid L-alanine [6,7]. These studies provided insights into the molecular conformation of the close surrounding of a nucleus but also indicated changes in their chemical environment.

To appropriately model the solute-solvent interactions in diverse environments, the theoretical models should be able to detail local effects associated with hydrogen bonds as well as long-distance effects. Significant deficiencies have been observed in current implicit solvent models, since they simplify the interplay between solute and solvent by describing the solvent with a simple dielectric constant [8]. An improvement of the description of local effects and hydrogen bonds may be achieved through the micro solvation model [9], in which the solute and a small number of solvent molecules are treated explicitly by quantum mechanics. Hence, local effects are represented by explicit solvent molecules, while long range effects are obtained by a dielectric constant. However, this approach has several pitfalls such as the selection of the number and positioning of solvent molecules may lead to artificial effects. In addition, it lacks the statistical characteristic of a liquid system. Recent developments in models combining quantum mechanics and molecular mechanics, namely QM/MM methods, have been more successful in describing the solute-solvent interaction and significantly improving the description of the properties compared to experiment [10–13]. One of the recent strategies for studying the solute in a solvent considers the free energy of the system as a function of the atomic solute coordinates. Under these circumstances the free energy gradient method combined with the QM/MM method has been proposed for the optimization of stable structures in the condensed phase. In such a model the normal modes and vibrational frequencies can be obtained in solution by the diagonalization of the Hessian matrix elements.

* Corresponding author.
E-mail address: valdemir.ludwig@ufjf.br (V. Ludwig).

https://doi.org/10.1016/j.molliq.2020.114109
0167-7322/

In this study, we investigated the solvation L-alanine in water solvent, making a systematic analysis of the stabilization of neutral and charged states. In the QM/MM method, we could have used any MM method such as Molecular Dynamics [14,15] (MD) or Monte Carlo [16–19] (MC) simulation techniques. In this work, we have favored the MC classical simulation due its easily implemented in our code, discussed below. We should note that, given the ergodic theorem, it is not expected that the structures generated by MD or MC, for the subsequent QM calculations, will differ much as long as they are statistically uncorrelated. The structure of the solute molecules in water is obtained by minimizing the free energy. The gradient of the free energy considered here is the gradient of the distribution of solute-solvent configurations sampled with the Boltzmann probability [20–22]. Thus, we present in detail the equilibrium structure of these neutral and charged alanine molecules in water. There are many studies of the structure of alanine in different environments and this study differs from those by an explicit consideration of the solvent in optimization process using the free energy structures gradient implementation in a QM/MM multiscale method. After that, important electronic proprieties such as NMR chemical shifts, magnetic shieldings and spin-spin coupling constants were determined by means of sophisticated quantum mechanical methods.

2. Methodology

First we considered the alanine molecule isolated and optimized the geometries using the second-order Moller-Plesset MP2/aug-cc-pVDZ level. An initial consideration of the solvent effects in the geometries was considered by adopting the Polarizable Continuum Model (PCM) [23] and again M2/aug-cc-pVDZ level. The structural and electronic properties of alanine molecule were investigated from theoretical point of view in gas phase and in aqueous solution at the MP2/aug-cc-pVDZ level, optimizing its rotamers and the zwitterionic structure. In addition, the charged species were also optimized in aqueous environment. Frequencies calculations ensured that these critical points correspond to true minima. More accurate structures are aimed by using a free energy gradient (FEG) consideration. In this case we have used the ASEC-FEG method [20,24]. This approach uses the sequential QM/MM (s-QM/MM) methodology [25] combining the Free Energy Gradient Method (FEG) [21,22] and an average solvent electrostatic embedding (EE) field (ASEC) [12] that represents the same electrostatic interaction obtained by the statistical average using all configurations sampled from the Monte Carlo simulation.

The ASEC-FEG approach allows to relax the geometry and simultaneously the atomic charges of the solute by adopting an iterative procedure of MC simulations and QM calculations. More details about the theory and computational process may be consulted in elsewhere [13,24].

The MC simulations were done employing the traditional Metropolis sampling method as implemented in the Dice package [26]. The alanine molecule was placed in a cubic box containing 1000 water molecules. Periodic boundaries conditions were applied by using the minimum image technique. The classical interaction was modeled by the Lennard-Jones (LJ) plus Coulomb potential. Initially, the LJ parameters for alanine were extracted from the all-atom optimized potential for liquid simulation parameters (OPLS-AA) [27], whereas Coulomb potential was described by a set of atomic charges obtained from an iterative procedure that brings the solute to an electrostatic equilibrium with the solvent. The CHELPG fit [28] is used at each step of the iteration employing the MP2/aug-cc-pVDZ level. These atomic charges were updated as well as the structure of alanine at each simulation step. For water molecules, the TIP3P [29] force field was adopted. The simulations were run in the NPT ensemble in normal conditions (P=1 bar and T=297.15 K) with a total number of 6×10^8 MC steps.

Using the energy autocorrelation function [17] at each optimization step, we select 400 solute-solvent configurations separated by less than 13% of statistical correlation to generated the ASEC mean field [12]. The forces and hessian matrix were calculated with respect to the alanine coordinates. With first and second derivatives, we are able to apply the Newton-Raphson technique to search the nearest stable configuration of the solute in solvated environment, although in the optimization procedure was adopted an approximate hessian.

The NMR calculations of the alanine molecule in water were carried out through the Gauge-Including Atomic Orbital (GIAO) [30] approach at the B3LYP/6–311++G(d,p) level of theory. The infrared spectrum was also computed via frequency calculations employing the same DFT functional and atomic basis set. In both calculations are required the evaluation of the full second-derivatives of the energy with respect to the atomic coordinates of the system. However, these hessian elements were estimated by taking into account just the atomic coordinates of the solute. All the QM calculations were performed with the Gaussian 09 program [31].

3. Results

3.1. Equilibrium geometry in water solvent

First, multiple minima of L-alanine were obtained in gas phase. It was identified 5 distinct stable structures, with all of them exhibiting positive frequencies. Illustration of the gas phase geometries of alanine are depicted in Fig. 1. In Table 1 are displayed the relative energies of these conformations. According to our results, the most stable structure is the AI_2, which is in agreement with previous works [32–35]. The next stable structure is the AIII, located 0.43 kcal/mol above the AI_2. It is worth mentioning that two rotations on the carboxylixc group is needed to reach the AIII from the AI_2, which means that either AII_1 or AII_2 will be accessed firstly. Moreover, no zwitterionic structure was determined, which is in accordance with our previous work about glycine molecule [36]. These findings show that the hypersurface of alanine is quite complex, containing several minima and transition states.

The minimum structure optimization of alanine in water started from the gas phase geometries structures obtained previously. In addition to optimizing the zwitterionic (Fig. 1a), also the anionic (Fig. 1b) and cationic (Fig. 1c) conformations were also optimized in aqueous environment. The zwitterionic structure has a net charge equal zero, but there is a migration of charge between the amino and carboxyl groups. This amplifies the polarization of the charge density, becoming it the most stable structure in aqueous solution. The A^{ZW} structure was stabilized by more than 78 kcal/mol in the ASEC-FEG method, while PCM model predicts that AIII structure is more stable than zwitterionic structure (around 1 kcal/mol above). This contrast between both methods could be related to the weakness of the PCM in dealing with specific interactions, for instance, hydrogen bonds. Our results with the ASEC-FEG method have the same tendency found previously by Takahashi et al. for glycine molecule [37]. These results indicate that the ASEC-FEG values are in better agreement with the recent experimental and theoretical work of Roggatz et al [4].

To obtain a better understanding into the minimum energy structures, it was examined the most relevant changes on the geometric parameters of alanine due to effects of water. All the bond distances and angles for the optimized structure in water may be consulted in table S1 of the Supplementary Material. In Fig. 2 is presented a schematic view of the C—C, C—N and C—O internal chemical bonds of the alanine molecule in gas and solvent phases. Our results employing the ASEC-FEG model were also compared with the experimental [38] data of neutron diffraction on crystal alanine samples in the neutral and zwitterionic forms. The zwitterionic structure optimized from the ASEC-FEG model always exhibits slightly higher values for all the distances in relation to experimental data, with the largest variation of 0.04 Å for C2-N3 bond. In the case of the PCM model, most distances were larger than the experimental data, but the largest discrepancy was of 0.025 Å for C2-N3 bond. It is important to mention that our

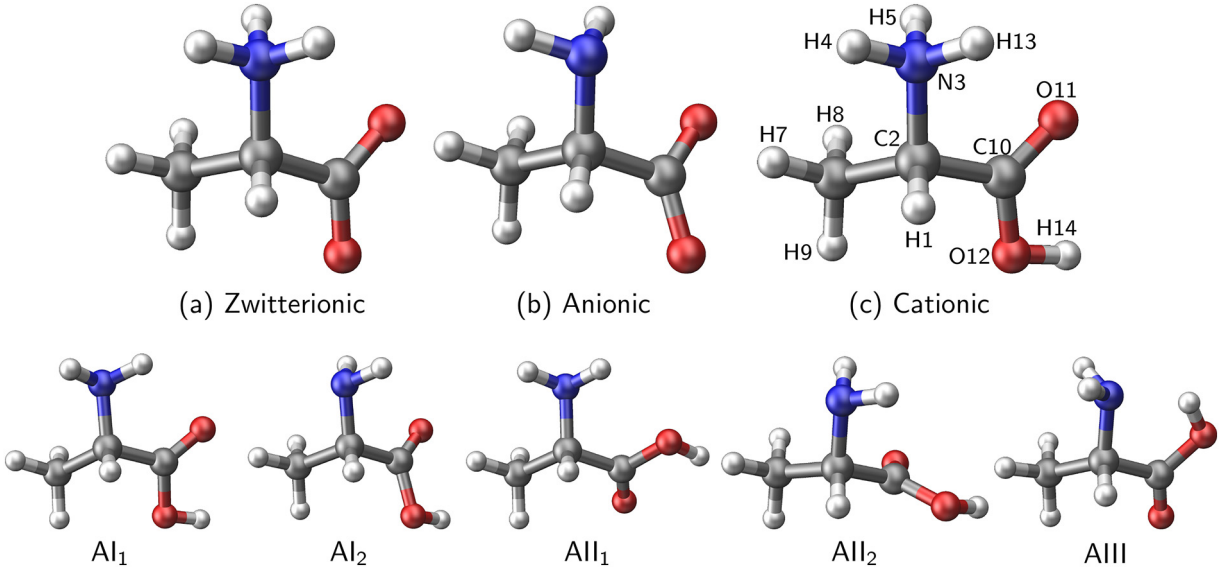

Fig. 1. The most stable optimized structure of alanine in its (a) zwiterrionic, (b) anionic and (c) cationic conformation in water solvent employing ASEC-FEG method. In addition, illustration of the gas phase structures of L-alanine. In both cases the MP2/aug-cc-pVDZ level was used for the optimization. Atom numbering is adopted in the text.

Table 1
Relative energies ΔE in kcal/mol for the optimized structures in gas and aqueous solvent phases. AI_2 was assumed as the reference structure.

$\Delta E[X-(AI_2)]$	AI_1	AI_2	AII_1	AII_2	AIII	A^{ZW}
Gas	1.46	0.00	1.81	1.10	0.43	–
Water (PCM)	0.97	0.00	1.55	1.08	−2.49	−1.38
Water (ASEC-FEG)	−4.33±1.38	0.00±1.44	−5.48±1.51	−0.57±1.51	3.01±1.19	−78.38±1.51

simulations and the experimental data were carried out in distinct environments.

Except for the C10-O12 bond, there is a stretch of the bond length from gas phase to water. For the charged optimized structures compared to ZW structure, the smallest variation were found for C—C bonds. On the other hand, the largest deviations appears on the C2-N3 for anionic molecule (difference of 0.035 Å) and C—O bond for cationic structure. These differences could be attributed to solute-solvent hydrogen bonds, which will be detailed below.

The average numbers of hydrogen bonds on the atomic sites of each molecule were estimated by analyzing the statistics of acceptor and donors bonds of the solute and solvent molecular structures. In Table 2 are displayed the average number of hydrogen bonds and their associated energies. As we can see, the A^{ZW} structure makes 9 hydrogen bonds with the solvent, in average, with approximately 6 on the carboxylic group and 3 on the amino (NH_3^+) group. For cation form A^+, the average number of hydrogen bonds drops to ~ 4.5, with less than 2 hydrogen bonds on the carboxylic group and around 3 on the amino group. The

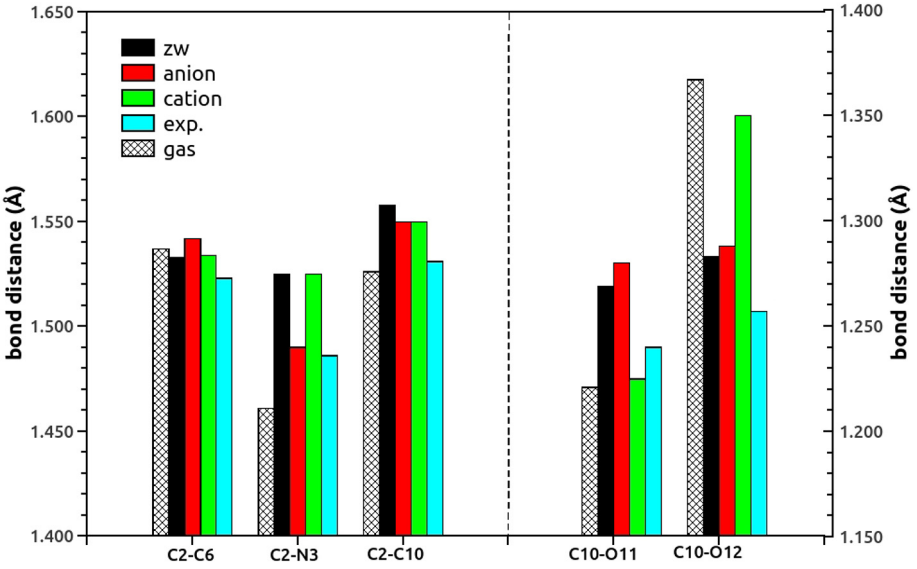

Fig. 2. Scheme of the calculated bond distances values for L-alanine molecule in gas and solvent phases at different protonation states using the ASEC-FEG model. Experimental data of neutron diffraction on crystal L-alanine samples in the neutral and zwitterionic form are exhibited for comparison [38].

Table 2
Statistical analyses of the average number of hydrogen bonds and their associated energies obtained from MC sampling.

Specie	H Bonds	⟨HB⟩	$\langle E\rangle_{HB}$
A^{ZW}	C=O ··· H–O	3.13	−28.58±4.68
	C–O ··· H–O	2.84	−25.98±3.93
	3×(N–H ··· O–H)	1.00	−41.67±3.67
	C–N ··· O–H	0.00	0.00
	sum	8.97	−96.23
A^+	C–O ··· O–H	0.16	−1.51±1.05
	C=O ··· H–O	0.60	−1.74±1.31
	O–H ··· O-H	1.00	−15.97±1.51
	3×(N–H ··· O–H)	1.00	−18.85±2.87
	C–N ··· O-H	0.00	0.00
	sum	4.46	−38.07
A^-	2×(C=O ··· O–H)	3.39	−15.53±2.57
	N–H_4 ··· O-H	0.99	−14.04±5.21
	N–H_5 ··· O–H	0.60	−3.54±2.57
	C–N ··· H–O	1.86	−9.61±3.66
	sum	10.83	−42.72

anionic alanine is the system which makes more solute-solvent hydrogen bonds, with 7 hydrogen bonds on the carboxylic group and almost 4 on the amine group. During the iterative procedure only the atomic charges of the solute are updated keeping the LJ parameters fixed throughout the optimization procedure. This may lead to some slight changes in the number of computed hydrogen bonds but a previous study by Martín and coworkers showed that the correction is not significant [39].

We also evaluated the energy of the hydrogen bonds. For the anionic form, with the largest number of hydrogen bonds on average, the energy due to these bonds is lower than those observed in the neutral form. In order to judge the most stable structure of alanine in water, we look the total average energy related to the hydrogen bonds: −96.23, −38.07 and − 42.72 kcal/mol respectively for A^{ZW}, A^+ and A^- states. Therefore, the cationic form A^+, with the less number of solute-solvent hydrogen bonds, was less stabilized in water solvent. It is interesting to note that the zwitterionic structure is the most stable, even though it is not the one with more hydrogen bonds. From the analyses of the average hydrogen binding energy and the number of bonds we found larger average energy for the A^{ZW}, of about 11 kcal/mol, for each hydrogen bond. For the charged structures this relative energy was lower. For the A^+ structure the medium energy was about 8.5 kcal/mol and for the A^- this energy reduces to 4 kcal/mol. As we can observe from these values, the highest binding energy value was obtained for the A^{ZW} and the lowest medium energy was calculated for A^- with the highest number of hydrogen bonds. Besides, the statistical analysis of hydrogen bonds on the amino and carboxyl groups of L-alanine reveals a decrease of 45% in the average hydrogen bond energy for the amino group of the A^+ state compared to the neutral situation (A^{ZW}). A pronounced decrease of 75% is calculated for the average hydrogen bond energy connected to the carboxyl group in the A^- state compared to the neutral case. Thereby, the effect of an acidic medium is to weaken the hydrogen bonds in the amine group, while the effect of a basic medium is to weaken the hydrogen bonds of the carboxyl group.

3.2. IR spectral characterization of the alanine conformers

In the ASEC-FEG methodology, effective equilibrium structures of the alanine molecule in solution were obtained and the normal modes and the vibrational frequencies in solution can be calculated. For the purpose to obtain the normal modes and vibrational frequencies of charged and neutral states of analine molecule in water solution, a vibrational study has been carried out for A^{ZW}, A^+ and A^- states. The pH-induced changes were estimated for the vibrational spectra. Experimental results for vibration spectra using diffuse reflectance infrared spectroscopy (DRIFT) [6,40] suggest strong intermolecular hydrogen bonds involving specific groups of L-alanine conformers at different pH scale. Likewise, experimental IR and Raman in solids and solution [41], and new sample method named DSD (dissolution, spray and deposition) [42] confirmed the role of solute-solvent hydrogen bonds in the IR spectra of L-alanine.

Theoretical and experimental data are displayed in Table 3 for the three states of L-alanine. A scaled factor of 0.96 [43] was adopted for

Table 3
Infrared frequencies of zwitterionic (A^{ZW}) and charged A^+ and A^- structures in aqueous solution.

Assignments	A^{ZW}			A^+		A^-	
	Exp.		ASEC	Exp. [6]	ASEC	Exp. [6]	ASEC
$\nu\,(OH)$	–	–	–	3109	3049.6	–	–
$\nu_{(ass)}\,(NH_2)$	–	–	–	–	–	3322	3327.7
$\nu_{(ass)}\,(NH_3)$	3072 [41]	3080 [6]	3095.9	3203	3027.5	–	–
			3013.3		2980.6		
$\nu_{(ass)}\,(CH_3)$	2991	2940	3064.9	2910	3067.8	2975	3037.2
			3042.7		3047.7		3031.4
$\nu_{(s)}\,(NH_3)$	3014	2988	2912	2979	2995.2	–	–
$\nu_{(s)}\,(NH_2)$	–	–	–	–	–	–	2984.4
$\nu\,(CH)$	2929		2997.2	–	2996.9	–	2984.4
$\nu_{(s)}\,(CH3)$		2884	2951.7		2957.5	2938	2937.4
$\nu_{(ass)}\,(OCO)$	1644	1620	1533.6	1720	1702	1581	1500.8
$\delta_{(sc)}\,(NH_2)$	–	–	–	–	–	1581	1606.3
$\delta_{(ass)}\,(NH_3)$		1590	1655.6	1602–1582	1654.4	–	–
			1648.1		1641.8		
$\delta_{(s)}\,(NH_3)$	1620	1519–1507	1555.5	1493	1557.4	–	–
$\delta_{(ass)}\,(CH_3)$	1455	1455	1426.2	1468	1431.8	1461	1431.8
			1429.7		1425.1		
$\delta_{(s)}\,(CH_3)$		1412	1361.7	1416	1350.3	1415	1330.1
$\delta\,(CH)$	1306	1307	1314.6	1310	1329.2	1305	1272.6
					1298.5		
$\delta(OH)$	–	–	–	1382–1346	1409.7	–	–
$\rho\,(NH_3)$	1150	1152	1201.5	1132	1148.9	–	–
$\rho\,(NH_2)\,(CC)$	–	–	–	–	–	1110–1071	1023.3
$\rho\,(NH_3)\,(CC)$	1114	1114	1143.2	1115	1197.8	–	–
$\rho\,(CH_3)\,(CN)$	1014	1015	1073.3	1005	1066.4	1023	983.9
$\delta\,(OCO)_{inplane}$	850	849	792.3			828	804.6
$\delta\,(OCO)_{outofplane}$	646	649	740.1	611	704.9	778	753.7
$\tau\,(NH_3)$	482	486	622.8		643.5		

all the vibrational modes in order to take indirect account of the anharmonic effects. For these calculations, we considered only the ASEC model, which means that no explicit water molecule is included in the Infrared frequencies calculations.

Our theoretical result for the O—H stretching mode in the cationic structure is assigned at 3050 cm^{-1}, which is in good agreement with the experimental value reported by Garcia et al. [6]. For the asymmetric amino stretching, large variations in the frequencies were found in the experimental data: 3072, 3203 and 3322 cm^{-1} for A^{ZW}, A^{+} and A^{-}, respectively. Except for A^{+}, our results are close to experiment (3052 and 3327 cm^{-1} for A^{ZW}, A^{-}, respectively). For A^{+}, our findings were down shifted calculated at 3000 cm^{-1}. A frequency of 2995 cm^{-1} was obtained for the symmetric amino stretching for A^{+}, which is little overestimated in relation to the available experimental data (2980 cm^{-1}).

For asymmetric CH_3 vibration, our calculated values are 3055, 3057 and 3035 cm^{-1} for A^{ZW}, A^{+} and A^{-} respectively, which are overestimated in relation to experimental data (2965, 2910 and 2975 cm^{-1}). For symmetric methyl group, we found an overestimated result for A^{ZW} and a very close value of 2937 cm^{-1} for A^{-} compared to the experimental value of 2938 cm^{-1}. At the asymmetric (OCO), our results are slightly below the experimental values but follow the variation for the three cases studied.

In lower frequencies, bending vibrations modes of the amino group were little overestimated, with the greatest discrepancy of only about 60 cm^{-1}. Vibrational modes related to the out of plane and in plane motion of the carboxylic group were obtained. A good accordance was achieved with the experiment. The worst results were achieved for motion outside of plane.

In summary, the ASEC model was able to obtain good insights for the vibrational frequencies, and these results could be improved by the inclusion of anharmonic effects. Unfortunately, there is not an easy form to insert explicit molecule from the theoretical point of view in the calculation by virtue of the existence of mix in the frequencies involving solute and solvent molecules. In addition, the IR spectrum is sensitive to changes in pH in the environment. We can mention the asymmetric stretching (OCO) vibration that suffers a red-shift in basic pH environment, in comparison with neutral pH, while a blue-shift is observed in acidic pH environments. Our model is able to describe these variations on the asymmetric (OCO) vibration mode. The number of hydrogen bonds in the carboxyl group, at least 7 for the (A^{-}) and less than 2 for the (A^{+}) state, can be used to explain these changes.

3.3. NMR Properties: shielding, chemical shift and spin-spin coupling of alanine in water solvent

The analysis of the NMR properties was focused on the most stable L-alanine structures in water, which encompass the A^{ZW}, A^{+} and A^{-} conformations. Shielding and coupling constants were computed using the structures sampled from the ASEC-FEG methodology in two distinct ways. In the first approach, the solvent is represented by the corresponding point charges in the atomic position employing an average electrostatic embedding (ASEC), and in the second one, some water molecules are added explicitly in the calculation, i.e., the quantum mechanical system is composed of the solute and some surrounding solvent molecules. In the latter approach is converged the statistic average of the property under consideration. In fact, we included here 12 explicit water molecules in our model, coined as ⟨12W + P. C.⟩ model. This number of explicit solvent molecules was defined based on the average number of solute-solvent hydrogen bonds, and the selection of these water molecules was carried out via minimum distance criterion [44]. For comparison, these properties were also computed using the PCM continuous solvent model.

The difference in the isotropic shielding between the neutral, zwitterionic and charged molecules was determined and the respective results are depicted in the Figs. 3 and 4. The positive and negative variations are associated with either increase or decrease of the property for charged alanine with respect to the A^{ZW}. For the anion case, as shown in Fig. 3, an increase in the shielding values were observed for the amino group, on the N3, H4 and H5 atoms. The same profile was also identified for O11 atom of the carboxylic group, but using explicit solvent molecules model it becomes more negative (~12 ppm). A negative sign of around 10 ppm was also observed for the C6, C10 and O12 atoms. Except for the O11 atom, the PCM model leads to a difference in the chemical shift with the same sign in relation to the approaches adopted here.

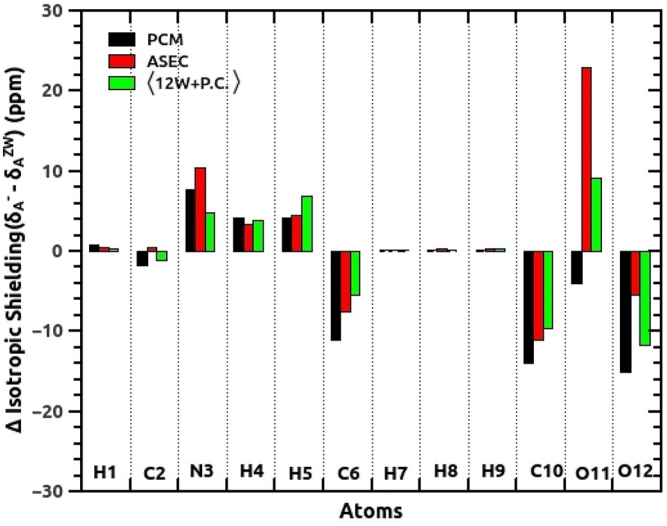

Fig. 3. Difference in the isotropic shielding values between the A^{ZW} and A^{-} structures.

For the cationic case, large variations on the oxygen atoms were induced by the protonation of the alanine molecule on the carboxylic group. A remarkable decrease of about 90 ppm in the shielding was seen for O11 atom, while the O12 atom exhibits opposite behavior. The PCM model gives the highest absolute values for shielding on the oxygen atoms. The insertion of 12 water molecules provides distinct trends, which could be related to organization of the closest solvent molecules around each oxygen atom. Besides, the deviation for other atoms is small (<10 ppm). Absolute values of the isotropic and anisotropic shielding can be seen in the Supplementary Material.

Chemical shifts were also evaluated and they can be compared directly with the experimental data employing reference standard values [5] for the atoms (H=31.83 ppm, N = − 180.16 ppm, C=183.45 ppm and O= 336.15 ppm related to DSS (Sodium trimethylsilylpropanesulfonate),

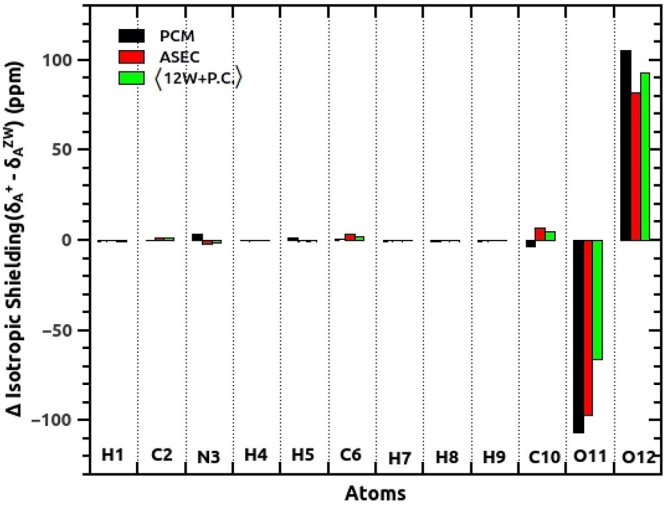

Fig. 4. Difference in the isotropic shielding values between the A^{ZW} and A^{+} structures.

nitromethane and water). Experimental data were extracted from measurements of di-alanine at room temperature taken from the reference [3]. Our theoretical results are showed in the Supplementary Material.

Since the chemical shifts are dependent on the choice of reference values, we focus our efforts in the differences of chemical shift of the charged and neutral forms. These results give pH-induced changes in chemical shifts and can be compared to experimental data as well as other previous theoretical works. As we can see in the experimental values [3], a negative sign is obtained for the hydrogen and nitrogen atoms for A^- molecule, whereas this behavior was observed in carbon and nitrogen atoms for A^+. Our theoretical results obtained from ASEC model reproduce satisfactorily the experimental trends for negative charged alanine, with high accuracy for hydrogen atoms. Explicit solvent molecules lead to a better agreement with the experiment, for instance, the electrostatic embedding ASEC wrongly predicts an negative sign on the carbon C2 atom, which is corrected with the ⟨12W + P. C.⟩ model.

For cationic alanine molecule, we reach good accordance with the experiment considering the solvent as point charges. A better agreement is achieved by means of inclusion of explicit water molecule. Depending on the atom, the PCM model also provides a good result, however, in some situations it is less accurate than the other approaches, it may even obtain an opposite sign regarding the experiment. Neither ASEC nor ⟨12W + P. C.⟩ methods are able to characterize the negative sign on the N3 atom.

For a general view, we compute the average absolute difference of the calculated and experimental data. They are presented in the last line of the Table 4. As we can see, our values significantly improve with the addition of explicit water molecules in the calculation, for which the absolute deviation drops from 1.81 ppm to 0.53 ppm for A^- molecule. On the other hand, similar amounts were achieved employing the three models for A^+, but a little bit smaller using ⟨12W + P. C.⟩ model. Dracínský and Bour predicted values of 0.44 ppm and 0.76 ppm respectively for A^+ and A^- molecules using Car-Parrinello molecular dynamics [46].

Additionally, spin-spin coupling constants for each equilibrium geometry A^{ZW}, A^+ and A^- respectively were also calculated and are available in Table 5. Previous results evidenced a pH dependence on the spin-spin coupling constants of alanine. A large conformational analysis has been done by Sychrovský et all [3] on the alanine dipeptide under pH changes. At the same study, they claim for improvements of the computational model, which is desirable by the reason of more accurate theoretical constants can drive to a better description of the peptide conformers. Our calculated results were directly compared to the experimental data as well as the mean absolute difference between theoretical and experimental data for each method. Considering all the possible spin-spin couplings of the alanine molecule, the mean absolute difference calculated with all approaches were less than 2 Hz. In regard to neutral zwitterionic system, the mean values obtained were 1.71 Hz and 1.80 Hz respectively for PCM and ⟨12W + P. C.⟩ models, while ASEC model gives the smallest value (1.1 Hz). For the anionic system, little bit better values were obtained by means of the ⟨12W + P. C.⟩ model, whereas the PCM model provides better values for A^+ system.

Table 5
One bond spin-spin coupling constants (Hz) for the PCM and the ASEC-FEG relaxed geometries for the zwitterionic A^{ZW} and the charged forms A^+ and A^-.

Specie	Spin-spin coupling pairs	PCM	ASEC	⟨12W + P. C.⟩	Exp. [3,45]
A^{ZW}	$C^\alpha - N$	−3.02	−3.71	−3.62	−5.7
	$C^\alpha - C^\beta$	34.37	33.90	32.93	34.9
	$C^\alpha - C$	50.33	56.03	55.13	54.0
	$C^\beta - H^\beta$	119.87	123.51	123.85	127.6
	$C^\alpha - H^\alpha$	146.70	142.39	140.56	145.1
	$C^\beta - H^\beta$	119.95	125.87	125.50	129.7
	$N - H^\alpha$	−1.37	−1.09	−0.77	0.0
	$N - C^\beta$	−0.45	−0.61	−0.42	0.0
	N-C	−0.06	0.02	−0.06	0.0
	$C^\alpha - H^\beta$	−3.10	−3.24	−3.25	−4.4
	$C^\beta - H^\alpha$	−2.97	−4.22	−4.25	−4.6
	$C - C^\beta$	−1.11	−1.14	−0.96	−1.2
	$C - H^\alpha$	−3.73	−4.26	−3.84	−5.0
	$N - H^\beta$	−3.27	−3.32	−2.70	−3.1
	$C - H^\beta$	3.59	4.11	4.40	4.2
	$H^\alpha - H^\beta$	6.46	6.64	7.00	7.3
	ΔJ	1.71	1.10	1.80	–
A^-	$C^\alpha - N$	−2.16	−1.84	−1.05	−4.3
	$C^\alpha - C^\beta$	30.58	30.72	29.79	35.2
	$C^\alpha - C$	53.97	55.90	53.68	52.7
	$C^\alpha - H^\alpha$	135.16	137.01	136.76	138.4
	$C^\beta - H^\beta$	119.87	123.51	123.85	127.6
	$N - H^\alpha$	−5.33	−4.76	−4.34	−2.2
	$N - C^\beta$	1.35	1.10	−0.63	0.0
	N-C	−1.12	−1.11	−1.07	0.0
	$C^\alpha - H^\beta$	−2.90	−3.09	−3.06	−4.3
	$C^\beta - H^\alpha$	−1.17	−2.83	−3.20	−4.7
	$C - C^\beta$	−1.11	−0.73	−0.62	0.0
	$C - H^\alpha$	−5.16	−5.14	−4.78	−4.3
	$N - H^\beta$	−0.87	−1.16	−1.00	−3.0
	$C - H^\beta$	3.63	4.11	4.41	4.3
	$H^\alpha - H^\beta$	6.53	6.16	6.48	7.3
	ΔJ	2.28	1.87	1.70	–
A^+	$C^\alpha - N$	−5.56	−5.67	−4.82	−6.6
	$C^\alpha - C^\beta$	33.49	33.29	33.20	34.1
	$C^\alpha - C$	61.48	62.38	60.14	59.6
	$C^\alpha - H^\alpha$	144.22	142.03	139.80	146.6
	$C^\beta - H^\beta$	125.85	127.08	125.89	131.0
	$N - H^\alpha$	−0.32	−0.19	−0.13	0.0
	$N - C^\beta$	−0.33	−0.51	−0.39	0.0
	N-C	−0.41	−0.35	−0.39	0.0
	$C^\alpha - H^\beta$	−3.42	−3.32	−3.37	−4.6
	$C^\beta - H^\alpha$	−4.05	−4.34	−4.23	−4.9
	$C - C^\beta$	−1.27	−1.15	−0.95	−1.3
	$C - H^\alpha$	−5.83	−5.12	−4.78	−6.0
	$N - H^\beta$	−3.46	−3.38	−2.70	−3.0
	$C - H^\beta$	3.63	4.11	4.41	4.6
	$H^\alpha - H^\beta$	6.66	6.67	6.48	7.3
	ΔJ	1.05	1.21	1.37	–

Table 4
Relative chemical shift in ppm of the charged forms. A^{ZW} structure was used as reference point.

Atoms	$\Delta\sigma(A^- - A^{ZW})$				$\Delta\sigma(A^+ - A^{ZW})$			
	PCM	ASEC	⟨12W + P. C.⟩	Exp. [3]	PCM	ASEC	⟨12W + P. C.⟩	Exp. [3]
H1	−0.72	−0.48	−0.31	−0.48	0.56	0.34	0.52	0.37
C2	1.84	−0.46	1.08	0.93	−0.30	−1.08	−1.13	−1.77
N3	−7.63	−10.50	−4.88	−6.60	−3.35	1.93	1.42	−2.20
H4	−4.09	−3.44	−3.86	–	0.34	0.01	0.29	–
H5	−4.10	−4.47	−6.84	–	−1.62	0.49	0.62	–
C6	11.11	7.54	5.40	4.25	−0.76	−3.30	−1.93	−0.83
H7	−0.12	−0.21	−0.09	−0.26	0.65	0.06	0.15	0.08
H8	−0.12	−0.24	−0.23	−0.26	0.48	0.14	0.23	0.08
H9	−0.10	−0.33	−0.34	−0.26	0.66	−0.09	−0.01	0.08
C10	13.97	11.05	9.69	8.94	3.26	−7.02	−4.48	−3.11
O11	3.96	−22.96	−9.12	–	106.85	97.27	66.15	–
O12	15.07	5.53	11.69	–	−105.22	−81.96	−92.68	–
$\Delta\sigma$	1.81	1.35	0.53	–	1.06	1.42	0.90	–

Overall, these results point out that models with explicit water molecules do not give much gains for the coupling constants. Same findings were reported in the study of NMR properties of water in liquid environment [20].

4. Conclusions

In the present work we performed a systematic study of the liquid water and pH induced polarization effects on L-alanine molecule. The investigation was performed by modeling the structure, NMR chemical shift and spin-spin coupling constants, and vibrational spectra. The structures were obtained by using a free energy gradient method implemented in a sequential QM/MM method. The detailed intermolecular hydrogen bonding effects described herein have important consequences on the stability of zwitterionic and protonated and deprotonated L-alanine structures. Our results predict that the structure of the A^{ZW} is more stabilized with almost 9 hydrogen bonds with an average binding energy of 10 kcal/mol. For the deprotonated A^{-} the average binding energy reduces to 4 kcal/mol. The local effects and hydrogen bonds were also very important in the NMR chemical shieldings and shifts. Values obtained with explicit water molecules reduce the calculated and experimental chemical shift differences and compared well with the experimental results. The far-ranging bulk solvent influence has main effect for the spin-spin coupling constants, a property in which the continuous solvent and average solvent models best describe compared to the explicit solvent models. In addition the results of the IR-spectra of the zwitterionic and protonated structures in water solvent show values with differences between the calculated and the experimental one of a maximum of 70 cm^{-1} for the region of the vibrational stretching modes. Although we have shown here only one application in estimating the stable ground state structure in aqueous solution, the present research protocol can be straightforwardly applied to processes in excited states allowing the study of photophysical properties.

CRediT authorship contribution statement

Valdemir Ludwig: Conceptualization, methodology, writing, review and editing. **Zélia M. da Costa Ludwig:** Writing, review and editing, project administration. **Danillo Valverde:** Software, validation, methodology, writing, review and editing. **Herbert Georg:** Software, supervision, reviewing. **Sylvio Canuto:** Review, supervision, project administration, funding acquisition.

Declaration of competing interest

The authors: Valdemir Ludwig, Zélia M. da Costa Ludwig, Danillo Valverde, Herbert Georg and Sylvio Canuto declare no conflict of interest in this manuscript.

Acknowledgement

The authors thanks to Brazilian funds CNPq, FAPESP, FAPEMIG and the high performance computer facilities of LNCC - Laboratório Nacional de Computação Científica. D. Valverde acknowledges financial support from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), under grant 2017/02612-4.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2020.114109.

References

[1] S.N. Timasheff, Protein-solvent interactions and protein conformation, Acc. Chem. Res. 3 (1970) 62–68.
[2] R.P. Rand, Probing the role of water in protein conformation and function, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 359 (2004) 1277–1284.
[3] V. Sychrovskỳ, M. Buděšínskỳ, L. Benda, V. Špirko, Z. Vokacova, J. Šebestík, P. Bouř, Dependence of the l-alanyl-l-alanine conformation on molecular charge determined from ab initio computations and NMR spectra, J. Phys. Chem. B 112 (2008) 1796–1805.
[4] C.C. Roggatz, M. Lorch, D.M. Benoit, Influence of solvent representation on nuclear shielding: calculations of protonation states of small biological molecules, J. Chem. Theory Comput. 14 (2018) 2684–2695.
[5] P. Bouř, M. Buděšínskỳ, V. Špirko, J. Kapitán, J. Šebestík, V. Sychrovskỳ, A complete set of NMR chemical shifts and spin-spin coupling constants for l-alanyl-l-alanine zwitterion and analysis of its conformational behavior, J. Am. Chem. Soc. 127 (2005) 17079–17089.
[6] A.R. Garcia, R.B. de Barros, J.P. Lourenço, L.M. Ilharco, The infrared spectrum of solid l-alanine: influence of pH-induced structural changes, J. Phys. Chem. A 112 (2008) 8280–8287.
[7] M.M. Quesada-Moreno, J.R. Avilés-Moreno, A.Á. Márquez-García, F. Partal-Ureña, J.J.L. González, L-serine in aqueous solutions at different pH: conformational preferences and vibrational spectra of cationic, anionic and zwitterionic species, J. Mol. 1046 (2013) 136–146.
[8] L. Onsager, Electric moments of molecules in liquids, J. Am. Chem. Soc. 58 (1936) 1486–1493.
[9] C.O. da Silva, B. Mennucci, T. Vreven, Combining microsolvation and polarizable continuum studies: new insights in the rotation mechanism of amides in water, J. Phys. Chem. A 107 (2003) 6630–6637.
[10] A. Muñoz Losa, I. Fdez. Galván, M.E. Martín, M.A. Aguilar, in: S. Canuto (Ed.), Solvation Effects on Molecules and Biomolecules: Computational Methods and Applications, Springer Netherlands, Dordrecht 2008, pp. 135–157.
[11] Hofer, T. S.; Randolf, B. R.; Rode, B. M. In Solvation Effects on Molecules and Biomolecules: Computational Methods and Applications; Canuto, S., Ed.; Springer Netherlands: Dordrecht, 2008; pp. 247–278.
[12] K. Coutinho, H. Georg, T. Fonseca, V. Ludwig, S. Canuto, An efficient statistically converged average configuration for solvent effects, Chem. Phys. Lett. 437 (2007) 148–152.
[13] H.C. Georg, T.S. Fernandes, N. Takenaka, K. Y., in: J. Leszczynsk (Ed.), Practical Aspects of Computational Chemistry III, Springer, New York 2014, pp. 231–248.
[14] C. Bistafa, Y. Kitamura, M.T.C. Martins-Costa, M. Nagaoka, M.F. Ruiz-Lopéz, Vibrational spectroscopy in solution through perturbative ab initio molecular dynamics simulations, J. Chem. Theory Comput. 15 (2019) 4615–4622.
[15] N. Takenaka, Y. Kitamura, M. Nagaoka, Efficient computational research protocol to survey free energy surface for solution chemical reaction in the QM/MM framework: the FEG-ER methodology and its application to isomerization reaction of Glycine in aqueous solution, J. Phys. Chem. B 120 (2016) 2001–2011.
[16] T. Malaspina, k. Coutinho, S. Canuto, Ab initio calculation of hydrogen bonds in liquids: a sequential Monte Carlo quantum mechanics study of pyridine in water, J. Chem. Phys. 117 (2002) 1692–1699.
[17] S. Canuto, K. Coutinho, From hydrogen bond to bulk: solvation analysis of the n-π* transition of formaldehyde in water, Int. J. Quantum Chem. 77 (2000) 192–198.
[18] S. Canuto, K. Coutinho, Solvent effects from a sequential Monte Carlo - quantum mechanical approach, Adv. Quantum Chem. 28 (1997) 90–107.
[19] K. Coutinho, C. Canuto, M.C. Zerner, A Monte Carlo-quantum mechanics study of the solvatochromic shifts of the lowest transition of benzene, J. Chem. Phys. 112 (2000) 9874–9880.
[20] H.C. Georg, S. Canuto, Electronic properties of water in liquid environment. A sequential QM/MM study using the free energy gradient method, J. Phys. Chem. B 116 (2012) 11247–11254.
[21] M. Nagaoka, Structure optimization of solute molecules via free energy gradient method, Korean Chem. Soc. 24 (2003) 805–808.
[22] N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, Transition-state optimization on free energy surface: toward solution chemical reaction ergodography, Int. J. Quantum Chem. 70 (1998) 95–103.
[23] S. Miertus, E. Scrocco, J. Tomasi, Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects, Chem. Phys. 55 (1981) 117–129.
[24] L.R. Franco, I. Brandão, T.L. Fonseca, H.C. Georg, Elucidating the structure of merocyanine dyes with the ASEC-FEG method. Phenol blue in solution, J. Chem. Phys. 145 (2016), 194301.
[25] K. Coutinho, R. Rivelino, H.C. Georg, S. Canuto, in: S. Canuto (Ed.), Solvation Effects on Molecules and Biomolecules: Computational Methods and Applications, Springer Netherlands, Dordrecht 2008, pp. 159–189.
[26] H.M. Cezar, S. Canuto, K. Coutinho, DICE: a Monte Carlo code for molecular simulation including the configurational Bias Monte Carlo method, J. Chem. Inf. Model. 60 (2020) 3472–3488.
[27] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (1996) 11225–11236.
[28] C.M. Breneman, K.B. Wiberg, Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis, J. Comput. Chem. 11 (1990) 361–373.
[29] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[30] K. Wolinski, J.F. Hinton, P. Pulay, Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations, J. Am. Chem. Soc. 112 (1990) 8251–8260.
[31] M.J. Frisch, et al., Gaussian 09 Revision E.01, Gaussian Inc, Wallingford CT, 2009.

[32] D.M. Upadhyay, A.K. Rai, D. Rai, A. Singh, A. Kumar, Ab-initio and density functional study of l-and d-forms of alanine and serine in gas phase and bulk aqueous medium, Spectrochim. Acta A 66 (2007) 909–918.
[33] B. Lambiea, R. Rameekers, G. Maes, On the contribution of intramolecular H-bonding entropy to the conformational stability of alanine conformations, Spectrochim. Acta A 59 (2003) 1387–1397.
[34] H. Zhanga, Z. Zhou, Y. Shi, Density functional theory study of the hydrogen-bonding interaction of 1:1 complexes of alanine with water, J. Phys. Chem. A 108 (2004) 6735–6743.
[35] P. Chaudhuri, S. Canuto, An ab initio study of the peptide bond formation between alanine and glycine: electron correlation effects on the structure and binding energy, J. Mol. Struct. (THEOCHEM) 577 (2002) 267–279.
[36] D. Valverde, Z.M. da Costa Ludwig, C.R. da Costa, V. Ludwig, H.C. Georg, Zwitterionization of glycine in water environment: stabilization mechanism and NMR spectral signatures, J. Chem. Phys. 148 (2018), 024305.
[37] Y.N.T. Takahashi, H. Kawashima, N. Matubayasi, A novel quantum mechanical/molecular mechanical approach to the free energy calculation for isomerization of glycine in aqueous solution, J. Chem. Phys. 123 (2005) 124504.
[38] M.S. Lehmann, T.F. Koetzle, W.C. Hamilton, Precision neutron diffraction structure determination of protein and nucleic acid components. I. The crystal and molecular structure of the amino acid l-alanine, J. Am. Chem. Soc. 94 (1972) 2657–2660.
[39] M. Martín, M. Aguilar, S. Chalmet, M. Ruiz-López, An iterative procedure to determine Lennard-Jones parameters for their use in quantum mechanics/molecular mechanics liquid state simulations, Chem. Phys. 284 (2002) 607–614.
[40] A.R. Garcia, R.B. de Barros, A. Fidalgo, L.M. Ilharco, Interactions of l-alanine with alumina as studied by vibrational spectroscopy, Langmuir 23 (2007) 10164–10175.
[41] B.Z. Chowdhry, T.J. Dines, S. Jabeen, R. Withnall, Vibrational spectra of α-alanine acids in the zwitterionic state in aqueous solution and the solid state: DFT calculations and the influence of hydrogen bonding, J. Phys. Chem. A 112 (2008) 10333–10347.
[42] X. Cao, G. Fischer, New infrared spectra and the tautomeric studies and purine and l-alanine with and innovative sampling technique, Spectrochim. Acta A 55 (1999) 2329–2342.
[43] R.D. Johnson, et al., NIST computational chemistry comparison and benchmark database, http://cccbdb.nist.gov/ 2019.
[44] H.C. Georg, K. Coutinho, S. Canuto, Solvent effects on the UV-visible absorption spectrum of benzophenone in water: a combined Monte Carlo quantum mechanics study including solute polarization, J. Chem. Phys. 126 (2007) 034507.
[45] M. Dračínský, J. Kaminský, P. Bouř, Relative importance of first and second derivatives of nuclear magnetic resonance chemical shifts and spin-spin coupling constants for vibrational averaging, J. Chem. Phys. 130 (2009), 094106.
[46] M. Dračínský, P. Bouř, Computational analysis of solvent effects in NMR spectroscopy, J. Chem. Theory Comput. 6 (2010) 288–299.