Computational chemical analysis of unconjugated bilirubin anions and insights into pKa values clarification

Esteban G. Vega-Hissi, Mario R. Estrada, Martín J. Lavecchia, and Reinaldo Pis Diez

Citation: J. Chem. Phys. **138**, 035101 (2013); doi: 10.1063/1.4773586
View online: http://dx.doi.org/10.1063/1.4773586
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v138/i3
Published by the American Institute of Physics.

Computational chemical analysis of unconjugated bilirubin anions and insights into pKa values clarification

Esteban G. Vega-Hissi,[1,a)] Mario R. Estrada,[1] Martín J. Lavecchia,[2] and Reinaldo Pis Diez[2]

[1]*Departamento de Química, Facultad de Química, Bioquímica y Farmacia, Universidad Nac. de San Luis, Chacabuco 917, (5700) San Luis, Argentina*

[2]*Centro de Química Inorgánica (CEQUINOR), Departamento de Química, Facultad de Ciencias Exactas, Universidad Nacional de La Plata - CONICET, calle 115 y 47, (1900) La Plata, Buenos Aires, Argentina*

(Received 22 October 2012; accepted 14 December 2012; published online 15 January 2013)

The pKa, the negative logarithm of the acid dissociation equilibrium constant, of the carboxylic acid groups of unconjugated bilirubin in water is a discussed issue because there are quite different experimental values reported. Using quantum mechanical calculations we have studied the conformational behavior of unconjugated bilirubin species (in gas phase and in solution modeled implicitly and explicitly) to provide evidence that may clarify pKa values because of its pathophysiological relevance. Our results show that rotation of carboxylate group, which is not restricted, settles it in a suitable place to establish stronger interactions that stabilizes the monoanion and the dianion to be properly solvated, demonstrating that the rationalization used to justify the high pKa values of unconjugated bilirubin is inappropriate. Furthermore, low unconjugated bilirubin (UCB) pKa values were estimated from a linear regression analysis. © *2013 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4773586]

I. INTRODUCTION

Unconjugated bilirubin (UCB), one of the end products of heme catabolism, is a tetrapyrrole dicarboxylic acid consisting of two dipyrrinone groups with propionic acid side chains. The crystal structure reveals that the molecule adopts a folded, "ridge-tile" conformation in which each propionic acid group makes three hydrogen bonds with the opposite dipyrrinone.[1] This structure was also found to be the most stable one through quantum mechanical calculations.[2] The fact that these groups are not available to interact with water was used to explain the lipophilicity and poor water solubility of UCB.[3]

In aqueous solutions UCB co-exists as three species in equilibrium,[4] the neutral diacid (H_2B), the monoanions (HB^-), the structure of which differs only in the position of lactamic ring subtituents, and the dianion ($B^=$), with different properties and functions such as diffusion of the diacid through lipid membranes, like blood brain barrier,[5,6] binding of the monoanion to ABC-transporters,[5] and binding of the dianion to human serum albumin.[7] Because the relative ratio of the three species depends on both, the pKa, the negative logarithm of the acid dissociation equilibrium constant, values of UCB and the pH of the solution, the accurate determination of pKa values is of pathophysiological relevance due to its effects: bilirubin encephalopathy in severely jaundiced neonates[5,8] and precipitation of calcium bilirubinate salts in gallstones,[9] among others.

The pKa of both carboxylic acid groups in water is subject of discussion as several experimental values are reported. McDonagh and Lightner support aqueous pKas values of ∼4.2 and ∼4.9[10,11] expected for aliphatic propionic acid groups (∼4.86). Ostrow and Mukerjee uphold pKas values of ∼8.1 and ∼8.4[12,13] providing a rationalization based on three theoretical factors: (a) the restricted rotation of the –COOH and $–COO^-$ group, (b) the effect of breaking the –OH |||| O = C< bond on ionization, and (c) the steric hindrance to solvation of $–COO^-$ group, each factor related to the intramolecular H-bonding of UCB.[13] Because of these theoretical factors are based on smaller molecules with structures largely different from UCB, in this article we review the proposed factors through quantum mechanical calculations on the whole UCB molecule to avoid invalid comparisons. Our results provide evidence that the alleged structural factors proposed[13] to account for the higher pKa values are inconsistent with a more rigorous analysis. Besides, low pKa values were estimated from our calculations.

II. COMPUTATIONAL METHODS

Initial UCB conformation was taken from previous studies.[2] All bilirubin anionic species were generated by subtracting a proton (or two) from the neutral species. We carried out geometry optimizations in the gas-phase at B3LYP/6-31+G(d)[14,15] and MP2/6-31G(d) levels of theory. Natural bond orbital (NBO) analysis was used to evaluate the donor–acceptor interactions in the optimized structures. Solvent effect was evaluated through geometry optimizations using the integral equation formalism polarizable continuum model (IEF-PCM) at B3LYP/6-31+G(d) level of theory. Because of continuum solvation models represent an approach of the averaged behavior of solvent molecules and although the models have been revised and improved,[16] our system requires the

a)Author to whom correspondence should be addressed. Electronic mail: egvega@unsl.edu.ar. Telephone: +54 (2652) 424027.

addition of explicit solvent molecules to overcome continuum models limitations as has been previously described[17,18] and to take into account the effect of solvent in the hydrogen bond network. Six explicit water molecules were placed around the hydrogen bond donor and acceptor atoms of the dipyrrinone and propionic (propionate) groups and the geometry of the resulting supermolecule was fully optimized at B3LYP/6-31+G(d) level of theory. UCB dianion was also optimized with 12 water molecules solvating both propionate groups. To further investigate the effect of water in the conformation of UCB species, ten spherical clusters were generated placing randomly 300 water molecules around each UCB species and semiempirical PM6 (Ref. 19) method including the recently developed DH+ hydrogen-bonding correction[20] was employed to optimize the geometry.

It is worth mentioning that in between the two limiting approaches to deal with solvation effects, that is, the simulation of the solvent through a sort of dielectric container, in which the solute is introduced on one hand, and the use of explicit solvent molecules interacting with the solute on the other, there is a proposal made by Warshel and coworkers to simulate solvent molecules by means of Langevin dipoles.[21,22] The use of Warshel's proposal could be a very interesting alternative to study bilirubin in solution and into a protein environment and its utilization in the future is being considered.

To explore further conformational changes of UCB species, molecular dynamics simulations were carried out within the canonical ensemble. The PM6-DH+ semiempirical method[20] was used as the force field in which nuclei move. Implicit solvent (water) effects were taken into account through the COSMO (conductor-like screening model) model.[23] The molecules were heated from 0 to 298 K in 0.1 ps. The temperature was then kept constant by coupling the system to a thermal bath with a bath relaxation time of 0.5 ps.[24] After an equilibration period of 5 ps, a 100 ps-long simulation was carried out. The time step for all the simulations was 1.0 fs. No periodic boundary conditions were used.

Further, estimation of solution phase pKa values of UCB was carried out following a modification of the procedure of Klamt *et al.*[25] Recognizing that purely theoretical pKa values strongly overestimate experimental ones, these authors propose to use a linear relationship to correct that deficiency,

$$\mathrm{pKa}_{calc} = A\Delta G_{diss}/2.303\ RT + B \tag{1}$$

in which the slope A and the intercept B are calculated from linear regression methods using a set of experimentally known pKa values, and the dissociation Gibbs free energy values, ΔG_{diss}, are calculated from the reaction models

$$\mathrm{AH_2}(aq) \rightarrow \mathrm{AH^-}(aq) + \mathrm{H^+}(aq), \tag{2}$$

$$\mathrm{AH^-}(aq) \rightarrow \mathrm{A^=}(aq) + \mathrm{H^+}(aq). \tag{3}$$

Since only approximate pKa values are of interest, absolute ΔG_{diss} values are not needed. Thus, the reaction models (2) and (3) are simplified as

$$\mathrm{AH_2}(aq) \rightarrow \mathrm{AH^-}(aq), \tag{4}$$

$$\mathrm{AH^-}(aq) \rightarrow \mathrm{A^=}(aq) \tag{5}$$

and the contribution of protons to ΔG_{diss} is taken into account implicitly into the parameters of the linear relationship. Then, ΔG_{diss} for first and second ionizations ($\Delta G_{diss}1$ and $\Delta G_{diss}2$, respectively) were calculated for the reaction models (4) and (5) as

$$\Delta G_{diss}1 = G_{aq}(\mathrm{AH^-}) - G_{aq}(\mathrm{AH_2}), \tag{6}$$

$$\Delta G_{diss}2 = G_{aq}(\mathrm{A^=}) - G_{aq}(\mathrm{AH^-}). \tag{7}$$

The free energy of solvated species, G_{aq}, was estimated according to

$$G_{aq}\,(\mathrm{X}) = G_{gas}\,(\mathrm{X}) + E_{aq}\,(\mathrm{X}) - E_{gas}\,(\mathrm{X})\,, \tag{8}$$

where G_{gas} is the free energy in gas phase, E_{gas} is the total electronic energy in gas phase and E_{aq} is the total electronic energy when solvent effects are accounted for. The additional term that takes into account the change from the gas phase standard state to the solution phase standard state was not necessary either because it is cancelled when free energies are subtracted.

A conformational search was performed on a set of seven dicarboxylic acids using the software Balloon.[26] Conformers were further optimized with the B3LYP/6-31+G(d,p) level of theory. Total electronic energies and gas phase free energies were obtained at the same level of theory. The polarizable continuum model with UFF atomic radii was used to simulate solvent (water) effects. The same strategy was followed for neutral bilirubin. The starting geometries of bilirubin monoanions and dianion are taken from the optimized geometries of the lowest-energy conformations of neutral and monoanion species, respectively.

All geometry optimized UCB species were subjected to frequencies calculations at the B3LYP/6-31+G(d,p) level of theory and the common frequency scaling factor of 0.964 for this level of theory was applied to obtain a better agreement with experimental data. Vibrational frequencies of γ-lactam, pyrrol and propionic acid were also computed at the same level of theory to obtain the vibrational stretching modes of the unperturbed (not hydrogen bonded) functional groups presented in UCB.

Density functional theory (DFT) calculations were done using GAUSSIAN 03.[27] The NBO 3.1 program,[28] implemented in GAUSSIAN 03 package, was used to analyze charge transfer stabilization energies associated with hydrogen bonds and lone-pair delocalization. All semiempirical PM6-DH+ calculations were done with MOPAC2009 software.[29]

III. RESULTS AND DISCUSSION

In this article we have focused our attention in the main conformational changes of UCB after the formation of anionic species (deprotonation) analyzing the geometric parameters of the optimized structures and intramolecular and intermolecular interactions, especially, hydrogen bonds. We also

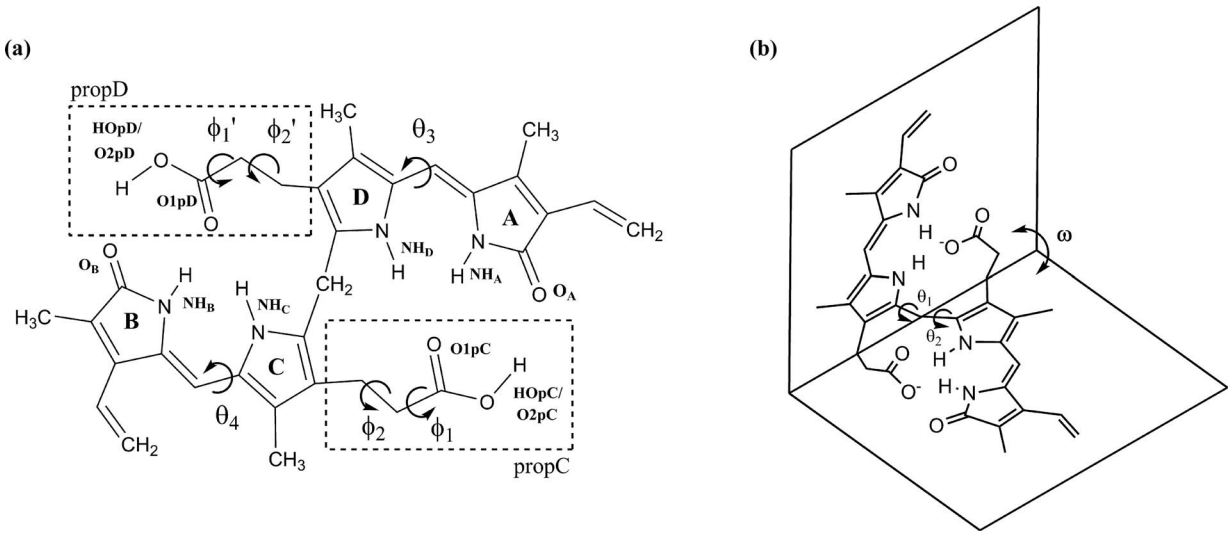

FIG. 1. Structure and definitions of UCB dihedral angles and labels of atoms and groups. (a) Schematic representation of the neutral specie; rings are labeled with capital letters. (b) Ridge tile conformation of the dianion.

performed an estimation of UCB pKa values. Figure 1 shows bilirubin diacid (panel A) and bilirubin dianion (panel B) with the dihedral definitions and atoms/groups labels which will be employed in the text.

A. *In vacuo* and implicit solvation

One of the major changes that occur after deprotonation is the rotation of carboxylate groups. Although the C=O group of the dissociating –COOH group is anchored by two H-bonds, the COO^- group rotates around the C–COO^- bond in both, gas, and aqueous (implicit solvation) phases. This can be seen by analyzing the deviation of ϕ_1 and ϕ'_1 dihedral values of the monoanions and the dianion from the corresponding value of the neutral species in Table I calculated at B3LYP/6-31+G(d) level of theory. Rotation also leads to shorter H-bonds distances between oxygen atoms of carboxylate group and the opposite dipirrinone.

Another interesting change when ionization takes place is the increase of the interdipyrrinone dihedral angle (ω) between the dipyrrinone groups which leads to a more extended conformation (Table I) being $B^=$ more extended than any of HB^- species. This is accompanied by increase of θ_1 and θ_2 dihedral values of both species.

It is important to notice that both monoanions present almost the same conformational behavior but in opposite halves of the molecule and that the deprotonation of **propC** or **propD** (defined in Figure 1) has minor effects on the other group because they are connected through a flexible backbone that attenuate any change. This may explain the small difference found between pKa_1 and pKa_2 values.

Parameter differences between gas phase and implicit solvent calculations are negligible. MP2/6-31G(d) calculations, which geometrical parameters are collected in Table SI of the supplementary material,[30] show the same trends as DFT method.

TABLE I. Hydrogen bond distances and geometrical parameters of UCB species. Distances are in Å and angles in degrees.

	B3LYP/6-31+G(d)				IEF-PCM/B3LYP/6-31+G(d)			
Parameters	H_2B	HB^- (propC)	HB^- (propD)	$B^=$	H_2B	HB^- (propC)	HB^- (propD)	$B^=$
O_A||||HOpC	1.660	...	1.646	...	1.663	...	1.650	...
NH_A||||O1pC	1.836	1.791	1.819	1.769	1.834	1.797	1.813	1.775
NH_D||||O1pC	1.952	1.659	1.961	1.683	1.950	1.691	1.942	1.704
O_B||||HOpD	1.652	1.640	...	...	1.654	1.640	...	...
NH_B||||O1pD	1.837	1.815	1.785	1.756	1.835	1.812	1.794	1.769
NH_C||||O1pD	1.948	1.958	1.665	1.683	1.943	1.940	1.694	1.704
θ_3	1.26	11.33	1.03	9.61	0.87	9.83	− 1.51	6.69
θ_4	1.39	0.37	10.03	8.86	0.93	− 2.05	8.83	5.67
ϕ_1	− 16.22	− 55.69	− 14.50	− 39.86	− 16.43	− 52.09	− 16.80	− 46.43
ϕ_1'	− 15.56	− 15.79	− 55.91	− 39.92	− 15.72	− 18.17	− 53.17	− 47.26
θ_1	61.87	64.28	69.31	77.14	62.00	65.52	69.36	75.73
θ_2	61.67	69.61	64.00	77.04	61.75	69.53	65.20	75.52
ω	95.56	106.88	106.59	128.15	95.60	107.22	107.14	124.61

TABLE II. Second-order interaction energies calculated at the MP2/6-31G(d) level of theory.

		$E^{(2)}$ (kcal/mol)			
UCB group	Interaction	H_2B	HB^- (propC)	HB^- (propD)	$B^=$
Ring A	$n_{OA(1)} - \sigma_{HOpC}{}^*$	9.86	…	9.83	…
	$n_{OA(2)} - \sigma_{HOpC}{}^*$	28.95	…	32.05	…
Ring B	$n_{OB(1)} - \sigma_{HOpD}{}^*$	9.76	9.81	…	…
	$n_{OB(2)} - \sigma_{HOpD}{}^*$	29.79	33.58	…	…
propC	$n_{O1pC(1)} - \sigma_{H}A^*$	6.29	1.91	7.49	2.16
	$n_{O1pC(2)} - \sigma_{NH}A^*$	10.13	23.48	11.39	29.62
	$n_{O1pC(3)} - \sigma_{NH}A^*$	…	2.89	…	0.95
	$n_{O1pC(1)} - \sigma_{NH}D^*$	5.38	6.47	5.06	8.14
	$n_{O1pC(2)} - \sigma_{NH}D^*$	…	24.45	…	14.74
	$n_{O1pC(3)} - \sigma_{NH}D^*$	…	16.53	…	22.44
propD	$n_{O1pD(1)} - \sigma_{NH}B^*$	6.49	7.68	1.95	2.21
	$n_{O1pD(2)} - \sigma_{NH}B^*$	10.08	11.33	23.85	30.44
	$n_{O1pD(3)} - \sigma_{NH}B^*$	…	…	2.88	0.90
	$n_{O1pD(1)} - \sigma_{NH}C^*$	5.03	4.78	6.47	8.22
	$n_{O1pD(2)} - \sigma_{NH}C^*$	…	…	23.81	14.03
	$n_{O1pD(3)} - \sigma_{NH}C^*$	…	…	16.30	21.98

B. NBO analysis

Second-order perturbative estimates of donor-acceptor (bond-antibond) interactions in the NBO basis (Table II) reveals that while an H-bond is broken at each ionization step losing two charge transfer interactions, some interactions are strengthened (higher energy values) and new interactions appear due to the reorientation (rotation mostly) of the carboxylate groups and the presence of a third lone pair over one oxygen of these groups. This result in a better stabilization of the anionic species supporting what have been suggested previously, that stabilization of monoanions of dicarboxylic acids leads to lower pKa1 values.[31] Monoanions show similar interaction energy values when comparing the equivalent interactions of deprotonated groups.

C. Infrared vibrational frequencies and hydrogen bonding energy estimation

Infrared (IR) spectrum of UCB and band assignments have been reported.[32–34] A sharp and intense band at 3410 cm^{-1} has been assigned to the N–H stretching mode expected due to the two N–H of the central pyrrole rings[32,34] which are intramolecularly hydrogen bonded to the opposite dipirrinone. This is evidenced that due to pyrrole, indole and similar compounds presents an N–H stretching frequency as high as 3490 cm^{-1}[35] and in the condensed phase hydrogen bonding occurs within amines causing a fall in the frequency.[36]

The band attributed to hydrogen bonded O–H group is observed as a very weak and broadband at 2559 cm^{-1} in the IR spectrum. The band at 1694 cm^{-1} is attributed to the carboxylic C=O stretching mode as being characteristic of molecules containing this group and the band at 1649 cm^{-1} is also assigned to C=O stretching mode. The frequency of carbonyl stretching vibration is slightly lowered down due to intramolecular hydrogen bonding.[34]

In different studies, the peak at 3260 cm^{-1} of the bilirubin IR spectrum has been assigned as the lactamic N–H stretching mode.[32,33] However, there are still ambiguous assignments regarding the lactamic N–H band of bilirubin that has not been discussed elsewhere. For instance, studies of Rai *et al.* have attributed the 3260 cm^{-1} band to the asymmetric OH stretching[34] without taking into account the previous assignments commented above.[32,33]

We have performed vibrational analysis in the harmonic approximation in order to compare the calculated vibrational frequencies of UCB against the available experimental data and the frequencies of characteristic groups (lactamic and pyrrolic N–H, and propionic acid O–H). The stretching frequencies of the groups involved in the hydrogen bonding network of UCB are presented in Table III. Computational calculations are useful to predict positions of the vibrational IR transitions and to help define these ambiguous assignments. For instance, our data reveal that the experimental band at 3260 cm^{-1} is more likely to correspond to the lactamic N–H stretching, supporting the assignments done by Yang *et al.* and Ferraro *et al.*[32,33] As can be seen in Table III (compare UCB values against free γ-lactam, pyrrole and propionic acid values),[35,37–39] hydrogen bonding in UCB causes a red-shift of all the stretching bands of the groups involved in the interaction.

The IR spectral measure of the hydrogen bond strength has been studied previously for a wide variety of

TABLE III. Calculated and experimental stretching frequencies (ν, in cm^{-1}) in UCB molecule, free γ-lactam pyrrole and propionic acid. Scaled values are denoted by a superscript SC.

	Computational calculations		Experimental data		
System	ν_{harm}	ν_{harm}^{SC}	ν_{exp}	Ref.	Assignments
UCB	3443	3319	3260	32	$\nu(N–H)_{lactamic}$
	3598	3468	3410	32, 34	$\nu(N–H)_{pyrrolic}$
	1730	1668	1649	34	$\nu(C=O)_{COOH}$
	2763	2663	2559	34	$\nu(O–H)_{propC}$
	2715	2617			$\nu(O–H)_{propD}$
γ-lactam	3669	3537	3455	37	ν(N–H)
	1790	1726	…	…	ν(C=O)
Pyrrole	3686	3553	3490	35	ν(N–H)
Propionic acid	3756	3613	3568	38, 39	ν(O–H)

TABLE IV. Enthalpies of hydrogen bonds (in kcal/mol) of UCB diacid species, calculated by Eq. (9) using experimental and computed vibrational shifts $\Delta\nu$ (cm^{-1}).

Hydrogen bond	$\Delta\nu_{exp}$	$\Delta\nu_{calc}$	$\Delta H_{\Delta\nu(exp)}$	$\Delta H_{\Delta\nu(calc)}$
O_A ‖‖HOpC	1009	950	24.22	22.80
O_B ‖‖HOpD		996		23.90
$NH_{lactamic}$ ‖‖O=C	195	218	4.68	5.23
$NH_{pyrrolic}$ ‖‖O=C	80	85	1.92	2.04

systems.[37,40,41] Hydrogen bonding produces a redshift and an intensification of the stretching ν(X–H) vibration. In order to estimate the hydrogen bond strength in UCB, we used the simple Badger-Bauer rule:[40]

$$-\Delta H\left(\text{kcal mol}^{-1}\right) = a\left(\left(\nu_0 - \nu\right)/\nu\right) \approx 0.024\left(\Delta\nu\right), \quad (9)$$

where a is a coefficient, ν_0 is the frequency of the unperturbed X-H vibrations, ν is the frequency of the perturbed vibrations (hydrogen bonded), and $\Delta\nu$ is the redshift of the stretching band.

The energy results collected in Table IV along with the hydrogen bond distances (see Table I) show that –NH ‖‖O=C< hydrogen bonds are weak, meanwhile –OH ‖‖ O =C < is stronger. However, when ionization takes place, the interaction of lactamic and pyrrolic N–H with the carboxylate group leads to a higher shift of the stretching frequencies towards values around 3027 cm^{-1} (data not shown). This results in the strengthening of both, lactamic and pyrrolic, –NH‖‖O=C< hydrogen bonds with energies values of 12.24 and 12.63 kcal/mol, respectively, stabilizing the anionic species.

It is important to note that although our frequencies values are somehow in disagreement with the experimentally determined values, applying the common frequency scaling factor of 0.964 for this level of theory a better agreement is obtained, taking into account that anharmonicity has not been considered. Because of hydrogen–bonded system execute highly anharmonic vibrations, appropriate methods or approaches such as the one described by Vianello *et al.*[42] should be required to improve the agreement of the calculated frequency values and its implementation in the future is being considered.

D. Molecular dynamics simulations

MD simulations show the conformational flexibility of bilirubin species. Rotation of COOH/COO^- can be appreciated following the temporal evolution of ϕ_1 and ϕ'_1 dihedrals in the simulations (Figure 2 and Figure S1 of the supplementary material[30]).

Propionic acid side chains conformational changes are more restricted in H_2B than in the anions due to the full H-bond network. However, the stability of this network cannot be assessed with implicit solvent models which cannot successfully characterize all solute–solvent interactions because they are unable to explicitly describe the hydrogen bond formation/dissociation between solute and solvent molecules.

Ionization of one propionic acid side chain (**propC**) leads to departure of ϕ_1 value from the value found in H_2B, while **propD** (protonated) preserves almost the same dihedral ϕ'_1 value (Figure 2(b)). The minimal deviation observed maybe due to a relaxation of the whole molecule.

The dianion shows the largest dihedral changes with ϕ_1 and ϕ'_1 values showing two populations (Figure 2(c)). Propionate oxygens (O1 and O2) of **propC** and **propD** side chains switch the interactions with NH groups of both pyrrolic and lactamic rings of the opposite dipyrrinone rotating dihedrals ϕ_1/ϕ'_1 (Figure 3(a)). Besides, these oxygen interactions also switch with subtle changes of dihedrals ϕ_2/ϕ'_2 (Figure 3(b)).

In all species dihedral angle ω fluctuates but folded structure remains. Flexibility allows the structure to open and close but limited by intramolecular steric interactions that tend to destabilize a wide spectrum of conformations including the helical, skewed, stretched, and linear shaped ones, and by the hydrogen bond matrix that stabilize the ridge-tile (folded) one as has been discussed previously.[43]

E. Explicit solvation (partially *ab initio* and fully semiempirical)

The geometry, charge distribution, and stability of intramolecular H-bonds can be affected by the presence of a proton acceptor and/or donor solvent because of the possible competition between intra- and inter-molecular hydrogen bonds.[44]

All geometrical parameters of optimized systems are collected in Tables V and VI. It is important to note that there is not a unique geometrical parameter to fully assess the conformational changes found in the studied UCB-water clusters. The combination of H-bonds distances and inter pyrrolic and lactamic rings dihedrals (θ_3 and θ_4) provide a good description of the solvent perturbation of H_2B.

We have found that interaction of water molecules with COOH group in H_2B leads to the distortion of the H-bond network and that weaker hydrogen bonds of the type –O ‖‖ HN– are more affected by the presence of solvent. The latter finding has been previously reported for neutral dimers using implicit solvation models[45] ascribing their results to the possible creation of optimal conditions for separate solvation of the donor and acceptor subsystems of the H-bond in competition with the geometrical requirements for the formation of this H-bond. When H_2B is locally solvated, DFT calculations show that while one intramolecular H-bond shortens, the other two elongate and θ_3 and θ_4 dihedral values change resulting in a complex with more interactions with water molecules (Figure 4(a)). Full solvation evaluated using the semiempirical method PM6-DH+, which predicts more accurately the energies and geometries involved in hydrogen bonding,[20] leads to the perturbation (elongation according to average values of the ten clusters) of all H-bonds, losing the planarity found in gas-phase geometries. In fact, this perturbation produces in some water clusters the rotation of COOH group (see ϕ_1/ϕ'_1 and ϕ_2/ϕ'_2 values of H_2B of clusters 7 and 8 in Table VI).

Distortion of the H-bond network of HB^- over the undissociated propionic acid group occurs in a similar degree as in

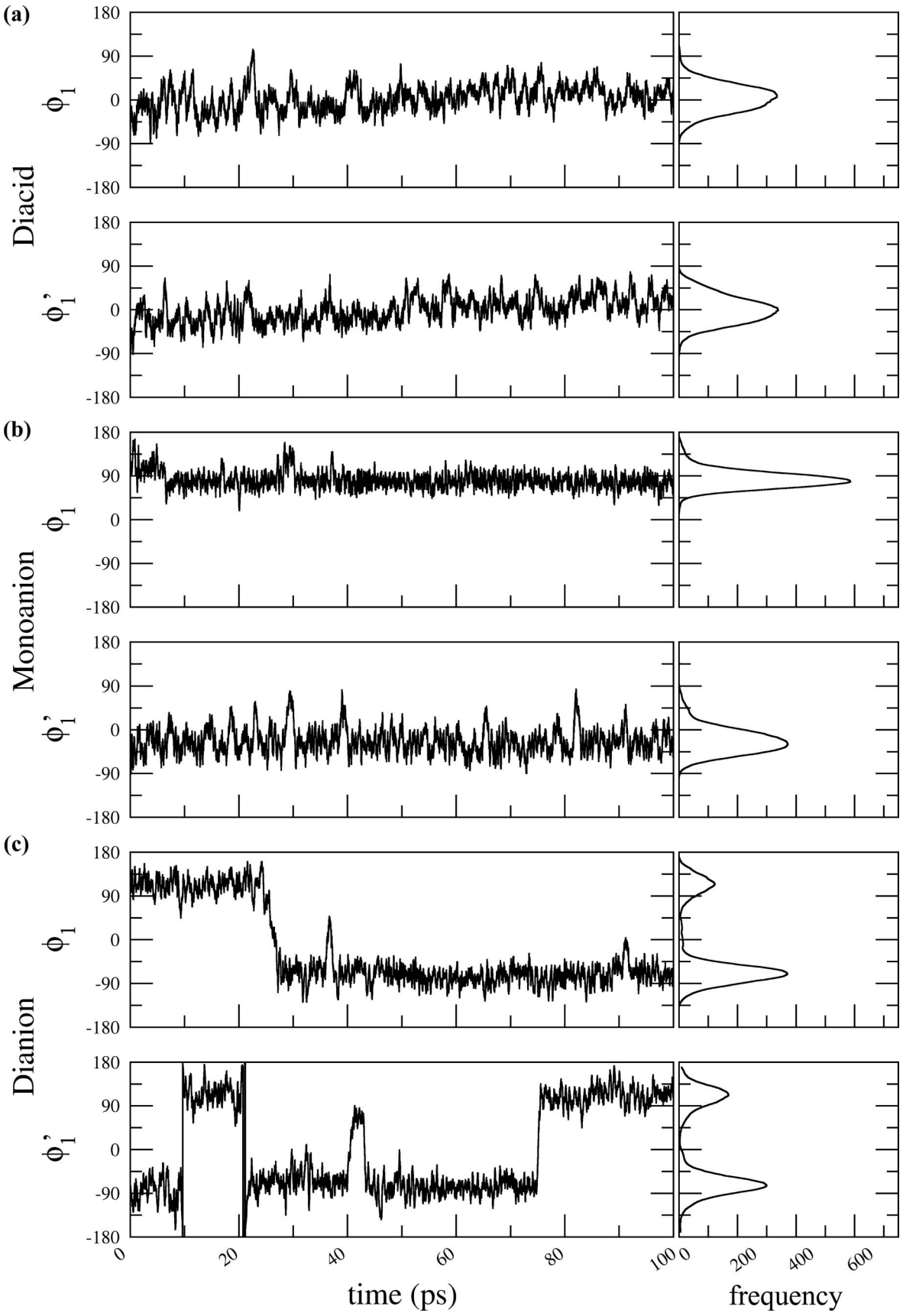

FIG. 2. Temporal evolution and frequencies of ϕ_1 and ϕ'_1 dihedrals of UCB species. (a) Diacid, fully protonated. (b) Monoanion deprotonated on propC. (c) Dianion. Angles are in degrees.

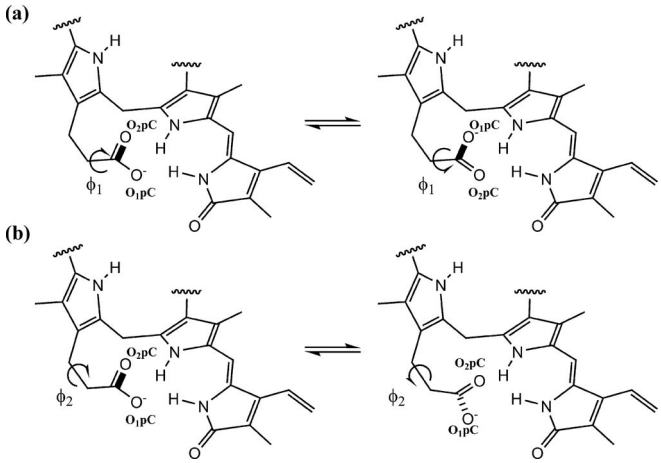

FIG. 3. Schematic representation of propionate C side chain rotations. (a) Rotation of ϕ_1. (b) Rotation of ϕ_2 exposes an oxygen atom of COO^- group above or below the opposite dipyrrinone plane.

H_2B. Rotation of COO^- groups of anionic species also occurs to different extent when explicit solvent is present (locally and fully solvated). This can be appreciated comparing ϕ_1/ϕ'_1 and ϕ_2/ϕ'_2 dihedral values of HB^- from Tables V and VI with the corresponding values of H_2B. Water molecules rearrange their position, giving rise to an intermolecular H-bond network between them and UCB anions (Figures 4(b) and 4(c)). This H-bond network prevents significantly the increase of the interdipyrrinone angle as can be seen comparing anionic species optimized in gas phase with the same species optimized in the presence of explicit solvent. Moreover, $B^=$ locally solvated in both propionates (12 water molecules) is less extended than the structure solvated only in one propionate (6 water molecules) and when UCB is fully solvated (water clusters), all species preserve almost the same interdipyrrinone angle similar to experimental values reported.[43,46]

The presence and persistence of folded, intramolecularly hydrogen-bonded conformation of bilirubin and its anions in solution is also experimentally supported by studies of Lightner and Nogales.[46] The hydrogen-carbon distances that have been determined from $^{13}C\{^1H\}$ heteronuclear Overhauser effects confirm that propionic acid carbonyl lies within hydrogen bonding distance to the opposite dipyrrinone lactam and pyrrole N–H groups as have been calculated in this study. Moreover, pyrrole and lactam ^1H NMR chemical shifts of UCB in $CDCl_3$ and DMSO-d6 (9.27, 10.69, 10.43, and 9.90 ppm, respectively)[11,47] account for the hydrogen bond effect (proton deshielding and a further downfield shift for these proton resonances) of the internal UCB hydrogen bond network.

TABLE V. Geometrical parameters of UCB species locally solvated obtained at B3LYP/6-31+G(d) level of theory. Monoanion is deprotonated on propC. Distances are in Å and angles in degrees.

Parameters	$H_2B.6H_2O$	$HB^-.6H_2O$	$B^=.6H_2O$	$B^=.12H_2O$
O_A||||HOpC	1.648	...	...	...
NH_A||||O1pC	2.965	1.832	1.839	1.861
NH_D||||O1pC	2.862	1.771	1.789	1.784
O_B||||HOpD	1.666	1.638		
NH_B||||O1pD	1.854	1.831	1.753	1.781
NH_C||||O1pD	2.037	1.955	1.689	1.733
θ_3	20.58	3.56	5.52	4.18
θ_4	7.91	0.82	5.02	4.91
ϕ_1	−16.41	−42.86	−60.63	−57.03
ϕ_1'	−18.64	−16.20	−60.13	−43.47
θ_1	51.35	63.42	77.60	70.30
θ_2	60.52	66.93	78.06	74.18
ω	91.94	103.34	126.60	116.15

F. Solvent accessibility

Hogle and Hammond analyzed the effects of alkyl substitutions in acetic acid concluding that bulky groups decrease the ionization constants because of reduced availability of the carboxyl oxygens for hydrogen bonding to solvent molecules.[48] In principle we might think that carboxylic and carboxylate groups of UCB, located in a crowded environment, would experience the same effect as acetic acid derivatives, i.e., raise its pKa. But when ionization takes place and –COO^- group rotates, one of the carboxyl oxygens protrude above (or below) the plane formed by the opposite dipyrrinone gaining access to solvent. Besides, lactam carbonyl group is more exposed to solvent too.

To inspect the steric hindrance to solvation, UCB-water clusters were analyzed counting water molecules surrounding the environment of carboxylic/carboxylate groups of the H-bond network (Figure S2 of supplementary material[30]) and averaging over the ten clusters. If the anionic UCB species has restricted access to solvent, less water molecules should be expected in the environment of these groups. Our calculations showed that there are more water molecules in close contact (within a radius of 3 Å, and farther too, see Figure 5) to polar groups of the H-bond network of HB^- (2.4) and $B^=$ (2.2) than H_2B (1.3). We think that the little difference between neutral and charged species may be because solvent perturbation of H-bonds leads to geometrical changes that allow better solvation of these groups in H_2B as has been proposed for other H-bonds networks.[45]

G. pKa estimation through linear regression

Figure 6 shows the linear regression obtained by fitting experimental pKa values of a set of dicarboxylic acids and calculated ΔG_{diss}. Calculated ΔG_{diss} values for the first and second ionizations of the selected acids are shown in Table SII of supplementary material.[30] The linear relationship obtained is

$$pKa_{calc} = 0.069\Delta G_{diss} - 14.76 \quad (10)$$

with a square correlation coefficient value of 0.834. Using Eq. (10) and ΔG_{diss} calculated for neutral, monoanionic, and dianionic bilirubin species, two sets of UCB pKa values were estimated considering that the monoanion could be formed by deprotonating either the propC group or the propD group, see Figure 1(a). When the monoanion is formed from the propD group, pKa1 and pKa2 values are 4.80 and 5.17, respectively,

TABLE VI. Geometrical parameters of UCB species water clusters optimized at PM6-DH+ level of theory. First data column presents gas-phase parameters. Distances are in Å and angles in degrees.

		$H_2B.300H_2O$									
Parameters	H_2B	1	2	3	4	5	6	7	8	9	10
O_A||||HOpC	1.682	2.407	1.699	1.737	1.773	1.636	1.770	2.633	1.713	1.849	1.664
NH_A||||O1pC	1.849	1.993	2.055	1.840	1.819	1.896	2.087	2.224	2.119	1.936	2.072
NH_D||||O1pC	1.973	2.065	2.029	2.122	1.898	2.077	2.074	2.127	2.132	1.977	2.049
O_B||||HOpD	1.688	1.961	1.872	1.990	1.769	1.750	1.776	2.052	2.119	1.873	1.841
NH_B||||O1pD	1.843	2.088	1.911	2.074	2.018	2.039	2.126	1.700	1.805	1.859	1.853
NH_C||||O1pD	1.978	1.975	2.011	2.003	2.034	2.208	2.081	1.921	1.904	2.049	1.983
θ_3	10.80	25.02	1.63	18.87	12.22	8.00	7.19	18.71	− 1.70	20.50	− 4.00
θ_4	12.02	4.07	9.05	4.12	11.18	− 1.71	5.12	− 3.41	− 5.17	13.02	9.47
ϕ_1	− 11.72	− 9.13	− 10.66	− 28.64	− 11.91	− 9.09	24.00	− 39.68	− 37.29	18.13	− 24.19
ϕ_1'	− 9.79	17.45	− 14.85	4.12	2.68	− 20.85	− 20.97	44.14	− 65.49	− 10.39	5.59
θ_1	56.57	57.94	57.88	63.63	61.34	59.46	58.05	64.54	61.36	52.51	57.18
θ_2	57.18	63.95	56.10	51.34	57.19	61.55	54.02	63.25	54.54	62.88	61.51
ω	93.50	97.80	89.08	93.12	96.64	97.23	91.48	95.25	86.10	95.74	92.40
		$HB^-.300H_2O$									
	HB^-	1	2	3	4	5	6	7	8	9	10
O_A||||HOpC	…	…	…	…	…	…	…	…	…	…	…
NH_A||||O1pC	1.667	1.809	1.802	1.778	2.638	1.835	1.791	1.889	1.804	1.847	1.863
NH_D||||O1pC	1.650	1.958	1.816	1.851	3.015	1.837	1.873	1.908	1.815	1.890	1.857
O_B||||HOpD	1.665	1.747	3.766	1.846	1.726	1.929	1.613	2.309	1.786	1.837	1.591
NH_B||||O1pD	1.820	1.943	2.520	1.853	2.063	1.710	1.858	1.887	1.844	2.238	1.992
NH_C||||O1pD	1.991	2.229	2.041	1.974	2.041	1.835	2.021	2.161	1.942	1.932	2.026
θ_3	18.78	9.46	6.33	22.75	5.08	3.67	15.53	− 14.20	− 3.11	23.56	− 7.38
θ_4	8.86	15.00	− 3.52	4.92	12.49	6.55	− 2.08	6.85	17.44	1.30	− 7.87
ϕ_1	− 61.25	− 64.97	− 21.60	− 29.17	− 39.20	− 63.32	− 61.88	− 11.05	− 67.18	40.36	− 41.77
ϕ_1'	− 10.99	3.71	48.36	− 15.45	− 18.81	− 71.72	− 16.02	− 34.97	− 18.62	− 14.49	− 33.46
θ_1	57.95	59.56	65.78	64.38	62.33	56.66	66.11	64.00	57.71	59.69	57.73
θ_2	62.77	64.76	60.97	62.60	58.21	61.99	70.18	59.91	60.94	55.45	64.20
ω	95.56	98.75	93.50	94.76	93.05	92.02	98.92	93.40	93.39	91.77	87.62
		$B^= .300H_2O$									
	$B^=$	1	2	3	4	5	6	7	8	9	10
O_A||||HOpC	…	…	…	…	…	…	…	…	…	…	…
NH_A||||O1pC	1.635	1.970	1.827	2.033	1.794	1.931	2.052	2.070	2.841[a]	1.800	1.758
NH_D||||O1pC	1.704	1.946	2.030	1.998	1.893	1.875	1.972	2.024	2.985[a]	2.061	1.820
O_B||||HOpD	…	…	…	…	…	…	…	…	…	…	…
NH_B||||O1pD	1.612	1.978	1.878	1.794	3.141[b]	1.833	3.126[c]	2.078	1.745	2.001	1.793
NH_C||||O1pD	1.725	1.866	1.816	1.886	3.512[b]	1.888	2.423[c]	2.063	1.963	2.359	1.960
θ_3	19.76	5.18	17.75	13.44	5.81	1.41	18.89	8.21	24.96	21.21	1.52
θ_4	18.85	7.61	8.33	16.46	14.85	6.16	4.20	9.51	2.51	16.10	12.84
ϕ_1	− 47.57	24.01	− 26.18	− 38.76	49.10	− 49.02	− 30.78	− 72.86	− 69.55	− 66.34	− 49.90
ϕ_1'	− 49.22	− 60.01	81.24	− 69.49	− 100.36	− 73.60	− 23.16	70.27	− 61.99	− 51.26	− 57.98
θ_1	67.86	59.71	59.56	57.27	57.73	57.85	61.59	53.19	52.54	58.70	65.36
θ_2	68.36	60.10	50.77	59.27	52.63	60.55	58.72	67.00	60.22	55.79	65.52
ω	103.92	93.34	84.33	90.84	87.24	89.88	94.72	92.73	91.90	98.49	99.9

[a]NH_A\|\|\|\|O2pC and NH_D\|\|\|\|O2pC distances: 1.815 and 1.941 respectively.
[b]NH_B\|\|\|\|O2pD and NH_C\|\|\|\|O2pD distances: 1.830 and 1.943 respectively.
[c]NH_B\|\|\|\|O2pD and NH_C\|\|\|\|O2pD distances: 1.949 and 2.694 respectively.

with a difference of 0.37 units. On the other hand, when the propC group becomes deprotonated to form the monoanion, both pKas are close to 5.00 and their difference is only 0.08 units. Both results are very close to the low experimental values reported and expected for aliphatic propionic acid groups, and are not greatly influenced by intramolecular hydrogen bonding. Moreover, present results suggest that the ionization process involves a monoanion formed after the propD group is deprotonated.

It has been recently informed the acid dissociation constants values of UCB obtained through computational methods.[49] While calculated pKa values for the first protonation step gave intermediate values between both experimental values reported, pKa values for the second protonation step were in severe disagreement with either experimental values. Taking into account that these values were based on a geometrical isomer of UCB without the characteristic intramolecular hydrogen bond network and that the methods employed were

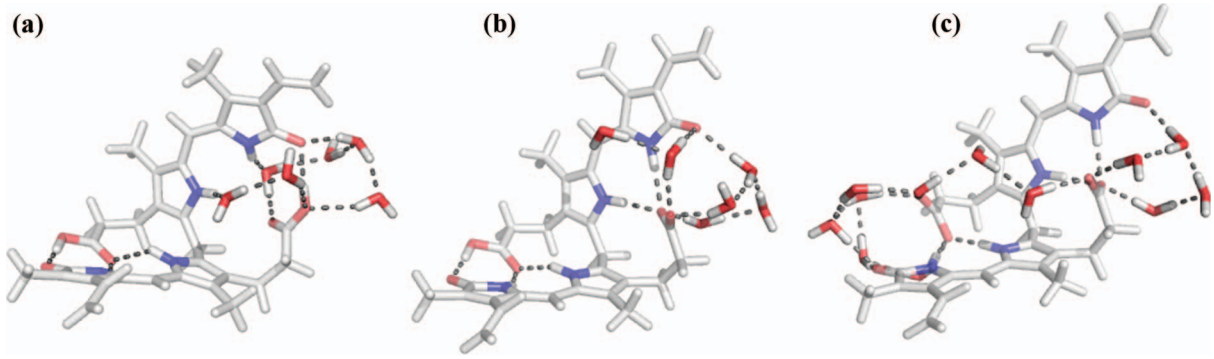

FIG. 4. Conformations of UCB species locally solvated. (a) Diacid ($H_2B.6H_2O$). (b) Monoanion ($HB^-.6H_2O$) deprotonated on propC. (c) Dianion ($B^=.12H_2O$).

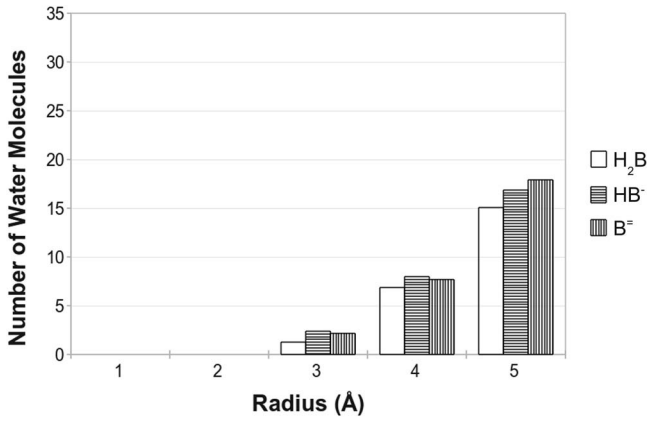

FIG. 5. Histogram of water molecules surrounding polar groups of UCB species.

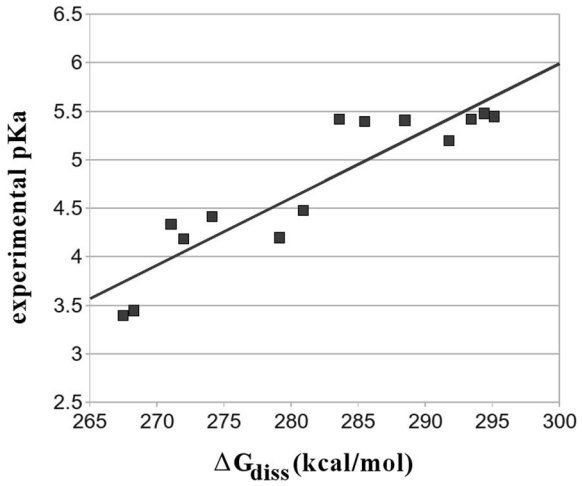

FIG. 6. Experimental pKa vs. calculated ΔG_{diss} values of a set of seven dicarboxylic acids. The linear regression is also shown.

validated with few (only three) monocarboxylic acids, these results could not be considered conclusive.

In contrast, our results were based on the biosynthetic, minimum energy isomer of UCB which presents the complex intramolecular hydrogen bonds and therefore, the effects of these interactions on the pKa estimation have been considered. Furthermore, the linear relationship was carried out using a set of dicarboxylic acids in order to account for both pKa values estimation.

IV. CONCLUSIONS

In the present study we have evaluated the factors related to the internal hydrogen bonding of UCB proposed by Ostrow and Mukerjee[13] that were used to justify rationally its high pKa values.

UCB is a flexible molecule,[2] however, ridge-tile structure is maintained in all studied species. The only major conformational change found in all analyzed environments (gas phase, implicit, and explicit solvent) after deprotonation is the rotation of $–COO^-$ groups which allow them to establish stronger interactions with the opposite dipyrrinone ring system, stabilizing the anionic species as was evidenced by NBO and vibrational frequency analysis. Explicit water molecules destabilize the intramolecular H-bond network in H_2B gaining access to set up a better solvation shell of UCB polar groups.

Although UCB present the $–COOH/–COO^-$ groups in a crowded microenvironment and establishing intramolecular hydrogen bonds that lead to think that rotation and solvation are difficult, our findings showed that: (a) rotation of $–COO^-$ group is not restricted; (b) new stabilizing interactions appear after an H-bond is cleaved at each ionization step; and (c) solvent accessibility is not inhibited. Due to the above, these factors are no longer credible/acceptable and pKa values rationalization based on them[13] may be misleading.

Moreover, both pKa values estimated in this work are consistent with the low experimental values reported.

ACKNOWLEDGMENTS

The authors thank CONICET, UNSL, and UNLP for the financial support. R.P.D. is member of the Scientific Research Career of CONICET. M.J.L. and E.G.V.H. are fellows of CONICET.

[1] R. Bonnett, J. E. Davies, and M. B. Hursthouse, Nature (London) **262**, 326 (1976).
[2] E. G. Vega Hissi, J. C. Garro Martínez, G. N. Zamarbide, M. R. Estrada, S. J. Knak Jensen, F. Tomás-vert, and I. G. Csizmadia, J. Mol. Struct.: THEOCHEM **911**, 24 (2009).
[3] A. F. McDonagh and D. A. Lightner, Pediatrics **75**, 443 (1985).
[4] J. D. Ostrow, P. Mukerjee, and C. Tiribellis, J. Lipid Res. **35**, 1715 (1994).
[5] J. D. Ostrow, L. Pascolo, D. Brites, and C. Tiribelli, Trends Mol. Med. **10**, 65 (2004).
[6] S. D. Zucker, W. Goessling, and A. G. Hoppin, J. Biol. Chem. **274**, 10852 (1999).
[7] J. Jacobsen and R. Brodersen, J. Biol. Chem. **258**, 6319 (1983).
[8] J. D. Ostrow, L. Pascolo, S. M. Shapiro, and C. Tiribelli, Eur. J. Clin. Invest. **33**, 988 (2003).
[9] D. Sutor and L. Wilkie, Clin. Sci. Mol. Med. **53**, 101 (1977).
[10] D. A. Lightner, D. L. Holmes, and A. F. Mcdonagh, J. Biol. Chem. **271**, 2397 (1996).
[11] S. E. Boiadjiev, K. Watters, S. Wolf, B. N. Lai, W. H. Welch, A. F. Mcdonagh, and D. A. Lightner, Biochemistry **43**, 15617 (2004).
[12] J.-S. Hahm, J. D. Ostrow, P. Mukerjee, and L. Celic, J. Lipid Res. **33**, 1123 (1992).
[13] J. D. Ostrow and P. Mukerjee, BMC Biochemistry **8**, 7 (2007).
[14] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
[15] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[16] J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).
[17] J. R. Pliego and J. M. Riveros, J. Phys. Chem. A **105**, 7241 (2001).
[18] E. F. da Silva, H. F. Svendsen, and K. M. Merz, J. Phys. Chem. A **113**, 6404 (2009).
[19] J. J. P. Stewart, J. Mol. Model. **13**, 1173 (2007).
[20] M. Korth, M. Pitonák, J. Rezaác, and P. Hobza, J. Chem. Theory Comput. **6**, 344 (2010).
[21] F. S. Lee, Z. T. Chu, and A. Warshel, J. Comput. Chem. **14**, 161 (1993).
[22] J. Florián and A. Warshel, J. Phys. Chem. B **101**, 5583 (1997).
[23] A. Klamt and G. Schüürmann, J. Chem. Soc., Perkin Trans. 2 **2**, 799 (1993).
[24] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[25] A. Klamt, F. Eckert, M. Diedenhofen, and M. E. Beck, J. Phys. Chem. A **107**, 9380 (2003).
[26] M. J. Vainio and M. S. Johnson, J. Chem. Inf. Model. **47**, 2462 (2007).
[27] M. J. Frisch, W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN03, Revision B.05. Gaussian, Inc., Wallingford, CT, 2003.
[28] E. D. Glendening, A. E. Reed, J. E. Carpenter, and F. Weinhold, NBO Version 3.1 (1993).
[29] J. J. P. Stewart, MOPAC2009, *Stewart Computational Chemistry* (Colorado Springs, CO, USA, 2008).
[30] See supplementary material at http://dx.doi.org/10.1063/1.4773586 for Figures S1 and S2 and Tables SI and SII.
[31] L. L. Mccoy, J. Am. Chem. Soc. **89**, 1673 (1967).
[32] B. Yang, B. Yang, R. Taylor, M. Morris, X. Wang, J. Wu, B. Yu, G. Xu, and R. Soloway, Spectrochim. Acta, Part A **49**, 1735 (1993).
[33] J. R. Ferraro, J.-G. Wu, R. D. Soloway, W.-H. Li, Y.-Z. Xu, D.-F. Xu, and G.-R. Shen, Appl. Spectrosc. **50**, 922 (1996).
[34] A. K. Rai, S. B. Rai, D. K. Rai, and V. B. Singh, Spectrochim. Acta, Part A **58**, 2145 (2002).
[35] R. A. Russell and H. W. Thompson, J. Chem. Soc. **1955**, 483.
[36] N. Fuson, M.-L. Josien, R. L. Powell, and E. Utterback, J. Chem. Phys. **20**, 145 (1952).
[37] M. V. Shablygin, D. N. Shigorin, and N. V. Mikhailov, J. Appl. Spectrosc. **3**, 40 (1965).
[38] J. Wenograd and R. A. Spurr, J. Am. Chem. Soc. **79**, 5844 (1957).
[39] E. M. S. Maçôas, L. Khriachtchev, M. Pettersson, R. Fausto, and M. Räsänen, J. Phys. Chem. A **109**, 3617 (2005).
[40] R. M. Badger and S. H. Bauer, J. Chem. Phys. **5**, 839 (1937).
[41] A. V. Iogansen, Spectrochim. Acta, Part A **55**, 1585 (1999).
[42] R. Vianello, B. Kovačević, G. Ambrožič, J. Mavri, and Z. B. Maksić, Chem. Phys. Lett. **400**, 117 (2004).
[43] D. A. Lightner, W. M. Donald Wijekoon, and M.-H. Zhangs, J. Biol. Chem. **263**, 16669 (1988).
[44] C. Soriano-correa, F. J. Olivares, A. Muñoz-Losa, I. F. Galván, M. E. Martín, and M. A. Aguilar, J. Phys. Chem. B **114**, 8961 (2010).
[45] J. A. Aquino, D. Tunega, G. Haberhauer, M. H. Gerzabek, and H. Lischka, J. Phys. Chem. A **106**, 1862 (2002).
[46] D. Nogales and D. A. Lightner, J. Biol. Chem. **270**, 73 (1995).
[47] J. O. Brower, D. A. Lightner, and A. F. McDonagh, Tetrahedron **57**, 7813 (2001).
[48] G. S. Hammond and D. H. Hogle, J. Am. Chem. Soc. **77**, 338 (1955).
[49] R. Borštnar, A. R. Choudhury, J. Stare, M. Novic, and J. Mavri, J. Mol. Struct.: THEOCHEM **947**, 76 (2010).