Theor Chem Acc (2009) 124:37–47
DOI 10.1007/s00214-009-0577-1

REGULAR ARTICLE

Effects of microsolvation and aqueous solvation on the tautomers of histidine: a computational study on energy, structure and IR spectrum

Amareshwar Kumar Rai · Weixin Fei · Zhiwen Lu · Zijing Lin

Received: 18 February 2009 / Accepted: 22 April 2009 / Published online: 17 May 2009

Abstract Microsolvation and combined microsolvation-continuum approaches are employed to investigate the structures and energies of canonical and zwitterionic histidine conformers. The effect of hydration on the relative conformational stability is examined. The strategy of exploring singly and doubly hydrated structures and the possible microsolvation patterns are described. We find that bonding water molecule may significantly change the relative conformational stabilities. In gas phase, the singly and doubly hydrated canonical forms are more stable than their zwitterionic counterparts. In solution, the continuum solvent model shows that bare zwitterionic form is more stable than bare canonical form by 1.1 kcal/mol. This energy separation is increased to 2.2 and 3.4 kcal/mol with inclusion of one and two explicit water molecules, respectively. We have also observed that the doubly hydrated structures obtained by combining two water molecules simultaneously to the solute molecule are preferred over the stepwise hydration. Hydrogen bond energies for the most stable hydrated histidine tautomers are determined by the atoms in molecules theory. The infrared (IR) spectra for the most stable singly and doubly hydrated structures of both histidine tautomers in gas phase are characterized. The stretching frequencies for NH of imidazole ring and OH of COOH are red shifted due to the hydrations. The IR spectra for the most stable zwitterionic tautomers in solution are also presented and discussed in connection with the comparison to the experiments. The p*K*a values obtained for the ring protonated zwitterions with two explicit water molecules appear to be in good agreement with the experiments.

Keywords Histidine–water complex · Histidine zwitterion–water complex · Conformers · DFT · CPCM

1 Introduction

It is well known that amino acids exist predominantly in their canonical form in gas phase while in solution they occur in zwitterionic form [1, 2]. So the exploration of the structures of an amino acid in both gas phase and in solution is an important issue of computational research. The interactions between amino acids and water molecules are of great interest because water is the natural medium for biological molecules and has important impact on their molecular structure [2–8]. Also, many gas phase processes involve microsolvation where water molecules affect the reactivity and structure of the biomolecules by electrostatic interactions [9]. So, it is equally important to characterize the process of hydration of amino acids in solution as well as in gas phase.

Two main approaches to include solvation are the dielectric continuum methods and the discrete microsolvation methods. Most of the solvation models used in the earlier work is based on the continuum approach. One of the most successful such model is the conductor-like polarizable continuum model (CPCM) [10–13] in which

Electronic supplementary material The online version of this article (doi:10.1007/s00214-009-0577-1) contains supplementary material, which is available to authorized users.

A. K. Rai · W. Fei · Z. Lin (✉)
Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, 230026 Hefei, Anhui, China
e-mail: zjlin@ustc.edu.cn

Z. Lu · Z. Lin
Department of Physics, University of Science and Technology of China, 230026 Hefei, Anhui, China

the interaction between the solute and the solvent is represented through a dielectric continuum. However, the continuum approach cannot describe the interaction between individual solute and solvent molecules which may be particularly important when hydrogen bonds are involved. The other approach is the cluster model or super molecular approach in which the water molecules are explicitly added at probable positions of solute molecule and the effects of microsolvation are examined as a function of the number of water molecules. The super molecular model is necessary for an accurate description of the near–neighbor interactions between the solvent and the solute but cannot satisfactorily account for the long-range electrostatic interactions between the solvent and solute. Therefore, the super molecular model should be combined with the continuum model to adequately characterize the behavior of amino acids in solution.

The hydration of several amino acids have been studied theoretically [14–22] and experimentally [5, 23]. However, the hydration and solvation properties of histidine still need to be investigated. Histidine is one of the most versatile protein residues. Its imidazole side chain has p*K*a near neutrality [24–26] and frequently serves either as an acid or a base in catalysis and in ligation of essential metal ions. This property is inherent neither in the standard nucleotide bases nor in the other natural amino acids. As a result, histidine is one of the residues most frequently used to form the active sites of protein enzymes. Consequently, the hydration and solvation properties of histidine are of special importance and should be carefully examined.

The conformational space of canonical histidine and their zwitterionic tautomers has been systematically explored in gas phase [27–29]. The present work examines the conformations of canonical histidine and their zwitterions with one and two explicit water molecules. To reliably determine the structures of histidine hydrated with one and two water molecules in gas phase, the possible hydration patterns are carefully examined and fully explored. The obtained conformers with different structural characteristics are examined with the continuum solvent model to determine their relative stabilities in aqueous solution. The hydrogen bond energies of the most stable structures of canonical histidine and its zwitterions hydrated with one and two water molecules are also presented. Besides, infrared (IR) spectra of hydrated histidine are analyzed.

2 Computational methods

2.1 Nomenclature

Figure 1 represents one of the tautomers of histidine i.e. His[N$^{\varepsilon}$H]. When hydrogen atom combines to the N_4^{δ} atom of the imidazole ring rather than the N_5^{ε} then it gives another tautomer of histidine named as His[N$^{\delta}$H]. For distinguishing the different conformers, each conformer is assigned with a numeral suffix to denote the ordering of stability of the bare conformers in gas phase.

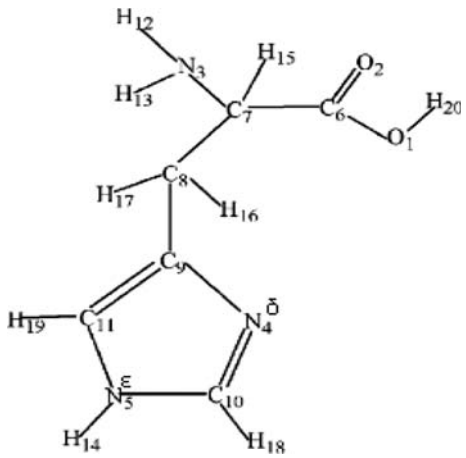

Fig. 1 Schematic structure of canonical histidine

When a proton transfers from the carboxyl to the amine group then zwitterionic tautomers are formed. A zwitterionic tautomer originated from canonical neutral histidine His[N$^{\varepsilon}$H] is denoted as zHis[N$^{\varepsilon}$H], while another zwitterionic tautomer originated from the other neutral tautomer His[N$^{\delta}$H] is denoted as zHis[N$^{\delta}$H], where "z" represents zwitterions. When the proton associated with carboxyl group of the neutral histidine transfers to the imidazole ring directly rather than to the amino group, it also forms a zwitterionic tautomer and is denoted as zHis[N$^{\varepsilon}$HN$^{\delta}$H]. Like the canonical tautomer, each zwitterionic conformer is assigned with a numeral suffix.

The hydrated histidine structures are labeled in the form of His[N$^{\varepsilon}$H]*m-i*, His[N$^{\delta}$H]*m-i*, zHis[N$^{\varepsilon}$H]*m-i*, zHis[N$^{\delta}$H]*m-i* or zHis[N$^{\varepsilon}$HN$^{\delta}$H]*m-i* to indicate the origins from which the hydrated clusters are formed. Here, *m* denotes the number of water molecules in the hydration cluster and *i* is the numeral suffix assigned to the bare conformer described above.

2.2 Computational method of exploring hydration

The stepwise hydration in gas phase has been carried out on the basis of stable histidine and its zwitterionic tautomers that have been systematically explored with ab initio calculations [27–29]. Water molecules are likely to attach at high polarity sites and form hydrated supermolecules. Several of the stable bare tautomeric conformers were selected and water molecules were added at all probable sites. Carboxyl and amino groups can easily combine with water molecules through hydrogen bonds. In addition, two nitrogen atoms of the imidazole ring of

histidine are also high polarity sites where hydrogen bonding with water molecules is possible.

The configurations of the bonded water molecules may vary and trial structures are produced accordingly. A single hydrogen bond structure may be formed by an electronegative atom on the amino acid molecule, i.e. oxygen or nitrogen, accepting a proton donation from the water molecule, or by the donation of a proton on the oxygen or the nitrogen atom to the water molecule, e.g., a hydrogen bond of water with NH of protonated imidazole ring. Optimization results, however, show that stable conformers favor multi-bond structures as shown in Fig. 2. So, trial structures with multi-bonds are considered for any two or more polar atoms where they might form hydrogen bonds with the water molecule as depicted schematically in Fig. 2. Note that the imidazole ring of histidine molecule has an exposed nitrogen and a protonated nitrogen that create complicated possible interactions with water molecules. Moreover, the protonation and deprotonation features of histidine are quite unique with both the deprotonation and the protonation processes occur on the imidazole ring [28].

When a second water molecule combines with the amino acid, the first water molecule behaves as an active moiety that could also interact with the second water molecule. Trial structures of the hydration are made by considering all the possible hydrate configurations. Figure 3 depicts some of the trial structures of neutral histidine with *cis* orientation of COOH. Our optimization results suggest that both water molecules prefer to stay near the carboxylic group (Scheme 1 of Fig. 3). The two water molecules can also hydrate at some other active sites on the amino acid as shown in Fig. 3. Similarly, trial structures were investigated with *trans* orientation of COOH of histidine tautomers where the most common stable structures are those where the water molecules bridge between the protonated and exposed nitrogen of the imidazole ring of histidine and the carboxylic group. Overall speaking, the larger the number of hydrogen bonds is, the more stable the structure is.

A complementary way to explore the structures of histidine hydrated with two water molecules is to consider trial structures formed by adding two water molecules simultaneously at probable sites of the bare histidine conformers. This process differs from the above stepwise hydration as in the latter the first water molecule may preclude some possible multi-water hydration forms. The complementary method is quite effective in producing the most stable conformers that may never be found in the process of stepwise hydration.

Trial structures for hydrated canonical and zwitterionic tautomers are made according to the above principles by starting with a large number of stable bare histidine conformers. All seven stable zwitterionic tautomers [29], two zHis[$N^{\varepsilon}H$] conformers, two zHis[$N^{\delta}H$] conformers and three zHis[$N^{\varepsilon}HN^{\delta}H$] conformers, are considered. There are a large number of the stable His[$N^{\varepsilon}H$] and His[$N^{\delta}H$] conformers and 20 of the most stable conformers for each tautomer are used for the trial structure generation. Numerous tests have been performed to validate that the

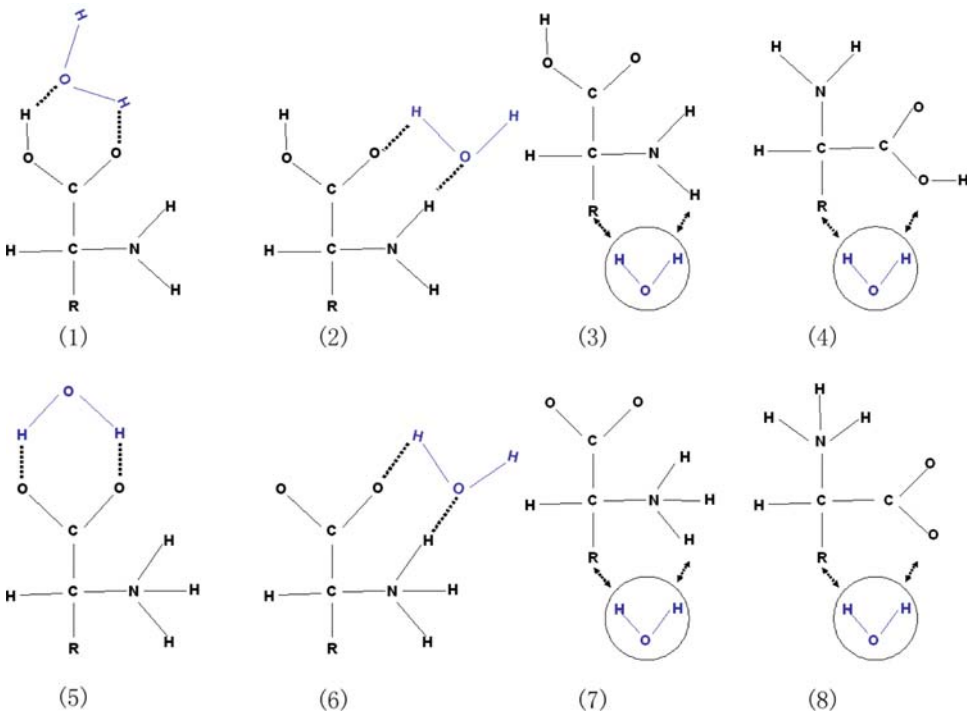

Fig. 2 Rough schematics of parts of configurations of hydrated amino acid with one water molecule. The actual situation is more complex due to the presence of various structures. *R* represents the side chain of amino acid. The *encircled H_2O group* indicates the possibility of multiple orientations

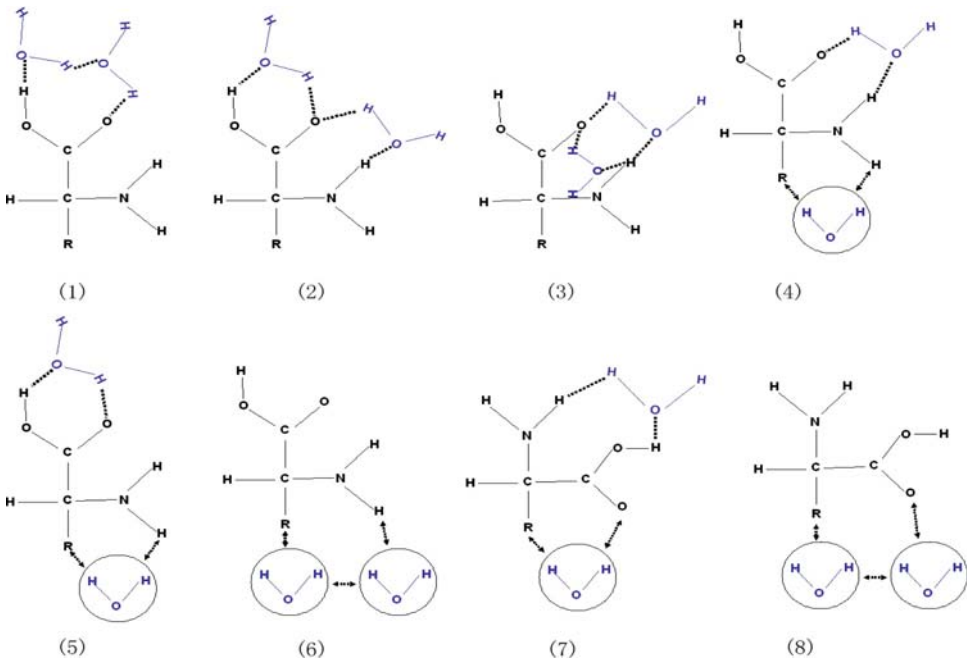

Fig. 3 Rough schematics for parts of the configurations of amino acid hydrated with two water molecules

use of other His[N$^{\varepsilon}$H] and His[N$^{\delta}$H] conformers in the trial structure generations does not produce new low energy hydrated supermolecules.

The hydrated structures were initially optimized at the B3LYP/6-311G(*d*) level of theory in gas phase using GAUSSIAN 98 program package [30]. The frequency calculations were performed for the optimized low energy structures to determine the true local minima. The frequencies were also used in the calculations of the zero point vibrational energies with a scaling factor of 0.96 as recommended by Scott and Radom [31]. The electronic energies were obtained by single point energy calculations at the B3LYP/6-311G(2*df*,*p*) and MP2/6-311G(2*df*,*p*) levels.

As the geometry optimization in solution hardly changes the structures optimized in gas phase [32], the solvent effect was simulated by performing a single point energy calculations using the CPCM self-consistent reaction field method (SCRF) with radii from the Universal Force Field (UFF) [33] at the B3LYP/6-311++G(*d*,*p*) level of theory. The use of the basis set of 6-311++G(*d*,*p*) instead of 6-311G(2*df*,*p*) for the solution calculations is to reduce the computational burden while maintaining the basis set adequacy for modeling the weakly bound complexes. The hydrogen bond energy was analyzed by the atoms in molecules (AIM) [34, 35] theory based on the potential energy density at the hydrogen bond critical point at the B3LYP/6-311++G(*d*,*p*) level.

The pK_a calculations have been carried out using the following relationship as discussed in detail in reference [36]:

$$\mathrm{p}K_a = [G_{gas}(A^-) - G_{gas}(AH) + \Delta G_{solv}(A^-) - \Delta G_{solv}(AH) - 269.0]/1.3644,$$

where A^- is the deprotonated counterpart of histidine species, *AH* (*AH* = bare, singly or doubly hydrated histidine). The subscript, "gas", denotes the gas phase. The free energy, *G*, of each state (either neutral or anion) is obtained by the single point electronic energy at the B3LYP/6-311++G** level and the thermal correction of free energy at the B3LYP/6-311G* level. The salvation free energy, ΔG_{solv}, is obtained from the CPCM calculation.

3 Result and discussion

3.1 Energy of bare and hydrated histidine

3.1.1 Energy of bare histidine with solvent model

Tautomers of bare canonical and zwitterionic histidine in gas phase have been systematically studied earlier [27–29] and the energies for their representative conformations are shown in Table 1. The energies of bare histidine tautomers with solvent model are also shown. Clearly, the relative stabilities of histidine conformers in solution are entirely different from that in gas phase due to the strong solvent effect. Even though the canonical conformers are much more stable than the zwitterions in gas phase, the most stable structures in solution are the zwitterionic tautomers of zHis[N$^{\varepsilon}$H] due to strong favorable electrostatic interactions with the solvent. The solvent effects on

Table 1 Relative conformational energies (kcal/mol) of histidine-$(H_2O)_m$ ($m = 0,1,2$) at the B3LYP/6-311G(2*df*,*p*) (B3LYP), MP2/6-311G(2*df*,*p*) (MP2) and B3LYP/6-311++G(*d*,*p*) with the CPCM model (CPCM) levels of theory

Tautomers	$m = 0$ [Bare]			Tautomers[a]	$m = 1$ [–H_2O]			Tautomers[a]	$m = 2$ [–$(H_2O)_2$]		
	B3LYP	MP2	CPCM		B3LYP	MP2	CPCM		B3LYP	MP2	CPCM
His[N^{ε}H]*m*-1	0	0	0	His[N^{ε}H]*m*-9	0	0	0	His[N^{ε}H]*m*-1	0	0.00	0.00
2	0.73	1.56	0.05	1	0.09	1.18	−1.26	1	0.85	1.63	−0.47
3	1.1	2.17	2.04	4	0.34	1.93	−4.12	4	1.68	4.61	−4.26
4	1.45	2.84	0.06	2	0.65	2.83	−3.79	9	1.81	3.68	−0.41
5	2.85	3.62	0.24	1	0.73	1.27	−2.83	1	2.11	4.09	−4.42
His[N^{δ}H]*m*-1	0.41	0.24	1.11	His[N^{δ}H]*m*-1	−0.53	−0.06	−2.5	His[N^{δ}H]*m*-11	0.42	0.07	0.87
2	1.54	2.75	1.24	7	0.5	1.59	−2.07	6	0.46	2.88	−2.60
3	2.54	2.04	3.61	2	0.59	1.77	−2.2	3	0.68	3.72	−3.95
4	2.99	3.09	4.49	3	0.73	1.97	−0.39	12	0.97	5.18	−5.30
5	3.36	4.25	4.32	12	0.88	3.04	−0.63	1	1.22	3.40	−3.41
zHis[N^{ε}H]*m*-1	11.48	11.4	−0.96	zHis[N^{ε}H]*m*-1	9.36	10.32	−4.59	zHis[N^{ε}H]*m*-1	6.7	5.98	−3.71
2	13.18	14.12	−1.11	2	10.04	12.2	−6.35	2	7.27	9.65	−7.42
				1	10.4	10.82	−5.2	2	7.38	10.23	−7.67
				1	10.41	9.65	−1.55	2	7.5	9.28	−5.22
				2	10.79	13.38	−6.32	2	7.89	10.36	−8.74
zHis[N^{δ}H]*m*-1	13.45	12.58	1.75	zHis[N^{δ}H]*m*-1	9.34	8.64	−1.31	zHis[N^{δ}H]*m*-1	7.68	7.10	−5.38
2	16.1	15.01	4.42	1	11.78	9.66	−0.59	1	9.34	10.34	−6.02
				1	12.36	13.08	−3.31	1	9.38	10.36	−5.98
				1	14.26	14.77	−2.82	1	9.41	8.20	−3.11
				2	14.89	13.8	0.48	1	9.67	8.44	−3.18
zHis[$N^{\varepsilon}HN^{\delta}$H]*m*-1	40.02	43.67	9.48	zHis[$N^{\varepsilon}HN^{\delta}$H]*m*-2	18.79	20.18	5.24	zHis[$N^{\varepsilon}HN^{\delta}$H]*m*-2	9.62	12.33	−3.06
2	41.58	44.97	9.79	1	19.03	18.65	5.66	1	10.73	11.73	−2.69
3	45.57	49.71	8.21	2	19.29	20.9	4.16	2	11.61	14.58	−2.90

The sums of the electronic energies and the zero point vibrational energies at the (B3LYP, MP2, and CPCM) level of theory are (−548.800352 a.u., −547.722267 a.u., −548.950779 a.u.); (−625.244451 a.u., −623.863269 a.u., −625.0430578 a.u.); (−701.693498 a.u., −700.164264 a.u., −701.9106708 a.u.) for His[N^{ε}H]-1, His[N^{ε}H]1-9 and His[N^{ε}H]2-1, respectively

[a] See the text for the nomenclatures. Same notation for different hydrated tautomers indicates that they are originated from the same bare conformers but the orientations of water are different

zHis[$N^{\varepsilon}HN^{\delta}$H] conformers are the most significant, though still insufficient to make the zHis[$N^{\varepsilon}HN^{\delta}$H] conformers as the most stable configurations. Note that the solvent generally favors the His[N^{ε}H] tautomers over the His[N^{δ}H] tautomers. This may be attributed to the larger dipole moments of the His[N^{ε}H] conformers (between 4.52 and 8.79 Debye) than that of the His[N^{δ}H] conformers (between 4.04 and 5.12 Debye).

3.1.2 *Energy of histidine hydrated with one water molecule*

The relative energies of up to five most stable structures of each histidine tautomer to which one H_2O is added are presented in Table 1. The most stable structure for each of the histidine tautomers attached with one water molecule is shown in Fig. 4. Some other low energy structures are included in the supporting document.

The hydration is seen not only to change the order of stability of different conformers but also to reduce the relative energy separation of the respective conformers. It is also seen that the mono hydrated structure formed from His[N^{δ}H] tautomer, His[N^{δ}H]1-1, is the most stable in the gas phase while His[N^{ε}H]0-1 is the most stable in the absence of H_2O. We also find that the global minimum conformer His[N^{ε}H]1-9 for His[N^{ε}H]1-*i* ($i = 1, 2,\ldots, 20$) derives from the 9th most stable His[N^{ε}H] conformer, indicating that the binding of a single water molecule significantly alters the order of stability of the bare tautomers. This result is similar to the cases for other hydrated amino acids [22, 32, 37]. As shown in Table 1, the dominant species of histidine hydrated with one water molecule are still the canonical ones.

When the solvent model is applied, the results show that the dominant species of histidine are zwitterions. The most stable zwitterions, zHis[N^{ε}H]1-2, is 2.2 kcal/mol more

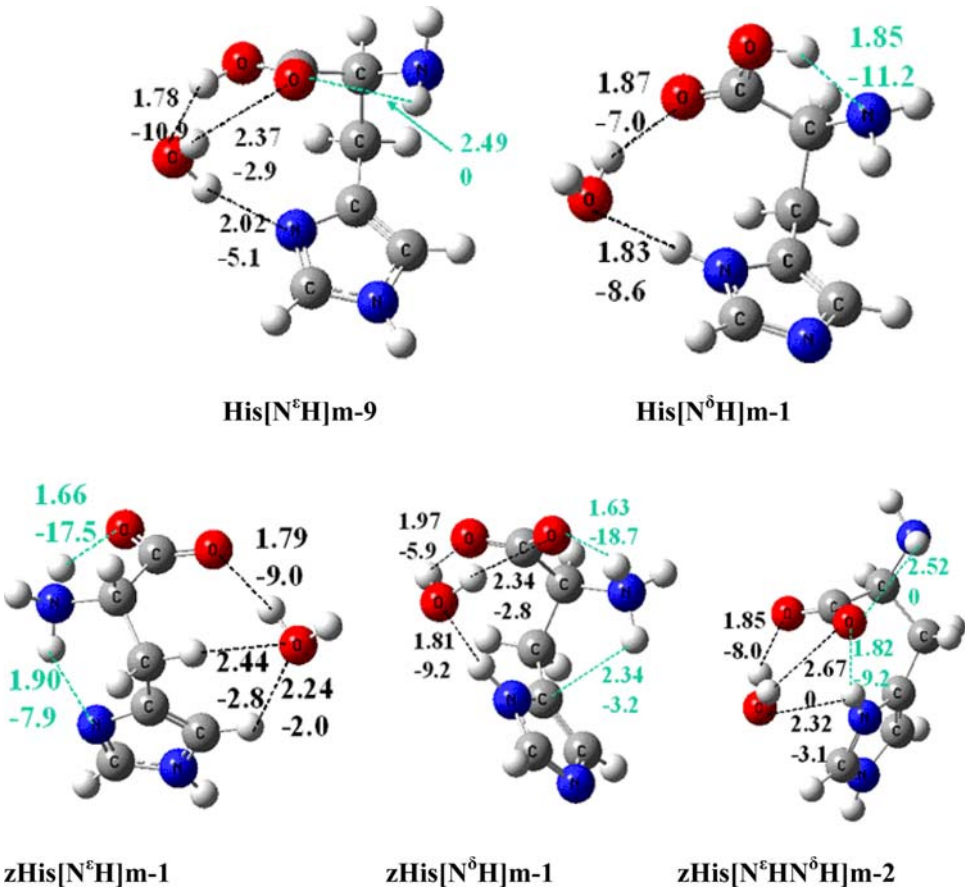

Fig. 4 The structures of the most stable conformers of histidine tautomers hydrated with one water molecule. Unit of the bond length (positive numerical value) in Å and bond energy (negative numerical value) in kcal/mol

stable than the most stable canonical structure, His[N$^{\varepsilon}$H]1-4. Notice that the corresponding energy gap with no explicit water molecule is only 1.1 kcal/mol. Adding explicit water molecules is necessary to demonstrate the predominance of zwitterions in solution, as also observed for other amino acids [22, 32, 37, 38].

Overall, the solvent model is more beneficial to the tautomers of His[N$^{\varepsilon}$H] hydrated with one H_2O than the His[N$^{\delta}$H] counterpart. This is similar to the case for bare histidine. For zwitterions with one H_2O, however, the solvent model benefits the zHis[N$^{\varepsilon}$H] and zHis[N$^{\delta}$H] tautomers more than the zHis[N$^{\varepsilon}$HN$^{\delta}$H] tautomers, a situation quite different from the bare zwitterions case.

3.1.3 *Energy of histidine hydrated with two water molecules*

Further information about the microsolvation effects is provided by examining the clusters of histidine with two water molecules. It is worth noting that the global minima for the tautomers of His[N$^{\varepsilon}$H], His[N$^{\delta}$H], zHis[N$^{\varepsilon}$H] and zHis[N$^{\delta}$H], His[N$^{\varepsilon}$H]*m*-1, His[N$^{\delta}$H]*m*-11, zHis[N$^{\varepsilon}$H]*m*-1 and zHis[N$^{\delta}$H]*m*-1, are found by adding the H_2O–H_2O cluster to the respective histidine tautomers, but are not found by the stepwise hydration method. The most stable structures of each tautomer of histidine with two water molecules are shown in Fig. 5. It is seen here again that the relative energy ordering of the hydrated conformers can be quite different from that of the bare conformers. For example, the global minimum of His[N$^{\delta}$H] tautomer hydrated with two water molecules, His[N$^{\delta}$H]2-11, is derived from the His[N$^{\delta}$H]-11 conformer with a relatively high energy of 4.91 kcal/mol above His[N$^{\delta}$H]-1. Also note that some hydrated tautomers with low energies are often derived from the same bare conformer and differ only in the orientation or the bonding site of the water molecules. For example, as shown in Table 1, the three hydrated tautomers of His[N$^{\varepsilon}$H] with relative energies of 0.00, 0.85 and 2.11 kcal/mol are all derived from the same lowest energy bare His[N$^{\varepsilon}$H] conformer. In the case of the first two hydrated tautomers, the two water molecules make a bridge between the carboxylic group and the NH of the imidazole ring with only a slight difference in orientations of the water molecules. In the third conformer, the water molecules make a bridge between amino group and an exposed nitrogen of the imidazole ring. The results show the preferential binding of water molecules toward carbonyl oxygen as compared to an exposed nitrogen of imidazole ring as well as the specific orientation of water molecules.

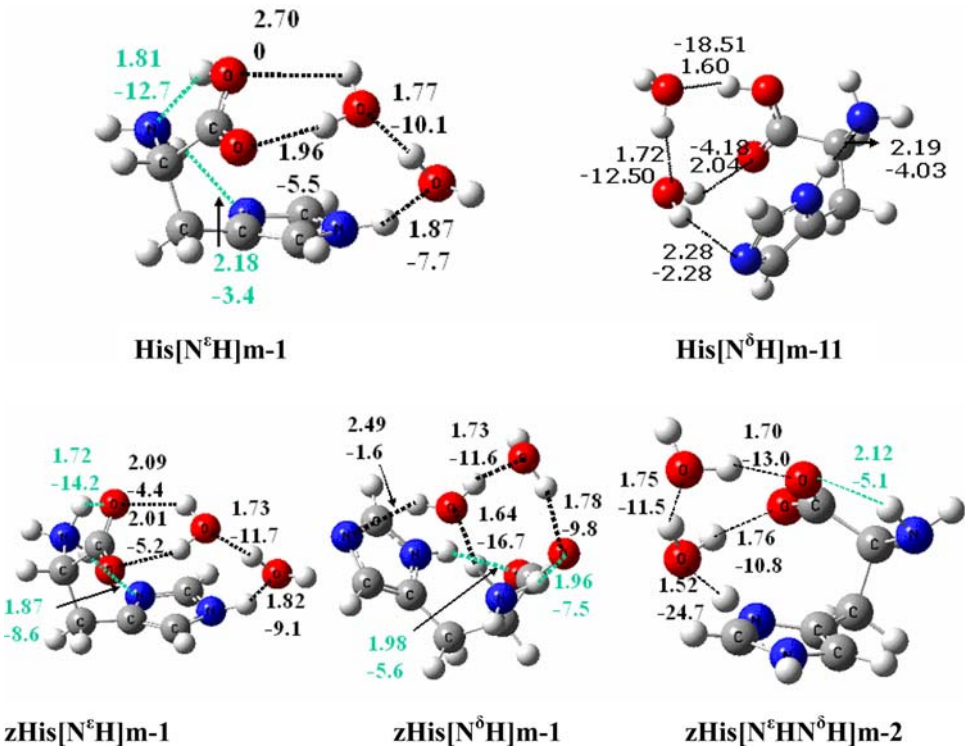

Fig. 5 The structures of the most stable conformers of histidine tautomers hydrated with two water molecules. Unit of the bond length (positive numerical value) in Å and bond energy (negative numerical value) in kcal/mol

It can be easily seen in Table 1 that zwitterionic histidine has stronger tendency to lower its energy on hydration. Moreover, the inclusion of more explicit water molecules inside the bulk solvent model is seen to increase the energy separation between the zwitterionic and canonical histidine structures. The results are very similar to the earlier studies on other amino acids [22, 32, 37–40]. Hydration with explicit water molecules is necessary for the study of the solvent effects on amino acids.

3.2 Structures of histidine

Figures 4 and 5 show the most stable structures for each histidine tautomer hydrated with one and two water molecules. The hydrogen bond lengths and energies are also presented based upon the AIM analysis. The hydrogen bond energy is determined using an empirical relationship $V_{cp} = 1/2E_{HB}$ [41] where V_{cp} is the potential energy density at critical points and E_{HB} is the gross hydrogen bond energy. The most stable conformer of His[N$^{\varepsilon}$H] tautomer with one water molecule is His[N$^{\varepsilon}$H]1-9 which is a folded conformation where the water molecule makes a bridge between the carboxyl group and an exposed nitrogen (N_4^{δ}) of the imidazole ring. The bond energies for the bonds OH$\cdots$OHH, $N_4^{\delta}\cdots$HOH and CO$\cdots$HOH are -10.9, -5.1 and -2.9 kcal/mol, respectively. The most stable conformer of His[N$^{\delta}$H] tautomer with one water molecule is His[N$^{\delta}$H]1-1 which is stabilized by a strong intra-molecular hydrogen bond as well as two inter-molecular hydrogen bonds.

With two attached water molecules, the most stable configurations of each canonical histidine tautomer, His[N$^{\varepsilon}$H]*m*-1 and His[N$^{\delta}$H]*m*-11, have comparable energies at both the B3LYP/6-311G(2*df*,*p*) and MP2/6-311G(2*df*,*p*) levels. The water–water hydrogen bond interaction in His[N$^{\delta}$H]*m*-11 is the strongest among the five tautomers shown in Fig. 5.

3.2.1 IR spectra of hydrated histidine in gas phase

In the gas phase, the most stable zwitterionic tautomers of singly and doubly hydrated histidine are of much higher energies than their canonical counterparts (Table 1) and are negligible in the equilibrium ensemble. However, the small energy differences between the most stable structures of each tautomer of canonical histidine with and without water molecules suggest that they may coexist in the gas phase. Also, the order of stability of the bare conformers is significantly changed on hydrations. In order to see the distinguishing features of the behavior of normal modes after hydration, the IR spectra of the most stable singly and doubly hydrated structures of each canonical tautomer are simulated and compared with the corresponding spectra of the bare tautomers (Fig. 6).

Infrared spectra in Fig. 6a show bands around 3,515 cm^{-1} corresponding to the N$^{\varepsilon}$-H stretching of the imidazole ring of the bare His[N$^{\varepsilon}$H] tautomer and the most stable singly hydrated His[N$^{\varepsilon}$H]*m*-9 tautomer. For the most stable doubly hydrated structure of His[N$^{\varepsilon}$H]*m*-1, the band is red shifted to

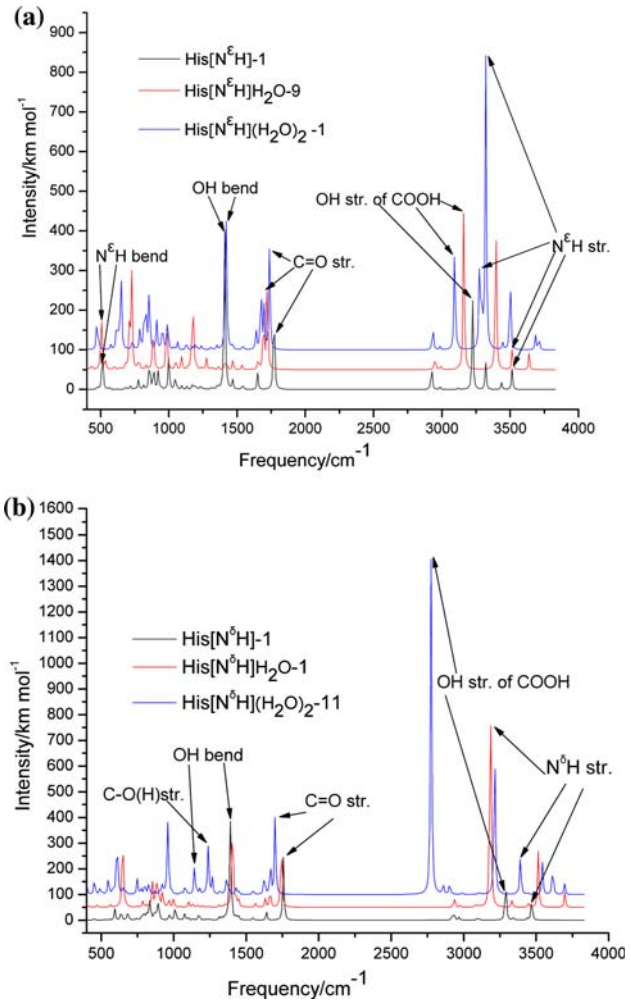

Fig. 6 IR spectra of the most stable conformers of His[N$^{\varepsilon}$H]-(H$_2$O)$_m$ and His[N$^{\delta}$H]-(H$_2$O)$_m$ (m = 0,1,2) in gas phase. **a** His[N$^{\varepsilon}$H]-(H$_2$O)$_m$, **b** His[N$^{\delta}$H]-(H$_2$O)$_m$

at 3,321 and 3,276 cm $^{-1}$ due to mixed vibrations of N$^{\varepsilon}$-H stretching and NH stretching of the amino group. The red shift is due to the intermolecular hydrogen bonding between N$^{\varepsilon}$-H of imidazole ring and water and between NH of the amino group and N$^{\varepsilon}$ of imidazole ring, as indicated in Fig. 5. The red shift is accompanied with enhanced intensity. The O–H stretching frequency is observed at 3,225 cm $^{-1}$ for the bare His[N$^{\varepsilon}$H] tautomer. The O–H stretching frequency is reduced in hydrated species due to the formation of the intermolecular hydrogen bond with water molecule and is at 3,094 cm $^{-1}$ in doubly hydrated structure of His[N$^{\varepsilon}$H]m-1. The red shift of the C=O stretching due to hydration is relatively small. The C=O stretching frequencies are 1,772, 1,722 and 1,739 cm $^{-1}$ for the bare, singly and doubly hydrated forms of His[N$^{\varepsilon}$H] tautomer, respectively. The OH bending frequencies are not much affected by hydration, but the intensity of OH bending in singly hydrated His[N$^{\varepsilon}$H]m-1 is greatly reduced.

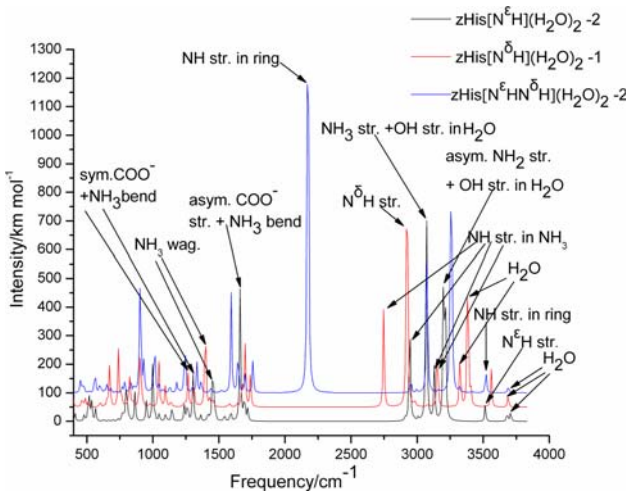

Fig. 7 IR spectra of the most stable zwitterionic tautomers of histidine in solution with two explicit water molecules

Compared with the cases for the His[N$^{\varepsilon}$H] tautomers, the red shifts of the N$^{\delta}$-H stretching and the OH stretching of COOH due to hydration are even more remarkable for the His[N$^{\delta}$H] tautomers (Fig. 6b). The red shifts of C=O stretching due to hydration are comparable for the His[N$^{\delta}$H] and His[N$^{\varepsilon}$H] tautomers. The OH bending vibrations of His[N$^{\delta}$H] tautomers are not much affected by hydration.

3.2.2 IR spectra of zwitterions

The IR spectra for the most stable zwitterionic tautomers of histidine with two explicit water molecules are given in Fig. 7 and additional details are available in the supporting document (Tables S1–S3). Experimental IR and Raman spectra of histidine at various pH are available in the literature [42, 43]. Comparison of the theoretical results with the experiments may be desirable, but one should be cautious about the suitability of the current models for direct comparison. The vibrational frequencies of polar groups are known to be strongly affected by the hydrogen bonds [44] and the inclusion of only two explicit water molecules is clearly insufficient for the description of the possible interactions between histidine and the solvent. Fortunately, the vibrational modes are sensitive only to the nearest neighbor interactions and the characteristic features of the vibrational spectra may be reflected in the histidine-(H$_2$O)$_2$ clusters with the corresponding characteristic local structures. That is, different structures of the histidine-(H$_2$O)$_2$ supermolecule may be required to identify different IR characteristic features observed experimentally. In other words, such comparison may be used to deduce the information about the structures of histidine in solution instead of proving the correctness of any single theoretical IR

spectrum. This assertion is illustrated with a brief discussion below.

Histidine is fully protonated at pH close to zero and the observed strong bands of C=O and C–O(H) at 1,737 and 1,259 cm $^{-1}$ [42] should be explained by the structures of canonical histidine with the carboxylic group surrounded by water molecules. A strong C=O band at 1,739 cm $^{-1}$ and a weak C–O(H) band at 1,192 cm $^{-1}$ are found with His[N$^{\varepsilon}$H]2-1 (Table S4), while a strong C=O band at 1,700 cm $^{-1}$ and a strong C–O(H) band at 1,239 cm $^{-1}$ are found with His[N$^{\delta}$H]2-11 (Table S5). Clearly, the C–O(H) band is described much better by the His[N$^{\delta}$H]2-11 configuration than by the His[N$^{\varepsilon}$H]2-1 configuration, indicating the carboxylic hydrogen should interact with H_2O as in His[N$^{\delta}$H]2-1 instead of amino N as in His[N$^{\varepsilon}$H]2-11 (Fig. 5). The result is quite natural since the amino N is protonated experimentally in this case.

As shown in Table 1, zHis[N$^{\varepsilon}$H]2-2 is the lowest energy structure in solution (but it is 7.89 kcal/mol above the global minimum in the gas phase by the B3LYP calculations) and may be expected to be representative of histidine at pH = 7.6. However, the bands of NH_3 bending and COO^- asymmetric stretching at 1,685 and 1,661 cm $^{-1}$ are found for this zHis[N$^{\varepsilon}$H]2-2, as compared to the experimental values of 1,646 and 1,600 cm $^{-1}$. The reasonable but less than satisfactory agreement may be attributed to the insufficient solvation of the NH_3 and COO^- groups in this zHis[N$^{\varepsilon}$H]2-2 (see Figure S2 for the structure). Properly solvated COO^- may be found in zHis[N$^{\varepsilon}$HN$^{\delta}$H]*m*-2 of Fig. 5 and the band of COO^- asymmetric stretching is found at 1,594 cm $^{-1}$ (Table S3), in good agreement with the experiment. Similarly, reasonable hydrogen bonding interactions for the NH_3 group are found in the zHis[N$^{\varepsilon}$H] 2-2 [7.38] structure of Figure S2 and the band of NH_3 bending for this structure is at 1,656 cm $^{-1}$ (Table S6), in satisfactory agreement with the experiment.

3.3 pK_a values

The pK_a values for each tautomer of histidine with 0, 1, and 2 explicit water molecules were determined based on only the respective most stable conformation. Each tautomer of histidine is characterized by two pK_a values due to two possible deprotonation sites and they are denoted as pK_{a1} and pK_{a2}. pK_{a1} corresponds to the deprotonation at the imidazole ring. pK_{a2} corresponds to the deprotonation at either the carboxylic group for His[N$^{\varepsilon}$H]*m-i*, His[N$^{\delta}$H]*m-i* or at the amino group for zHis[N$^{\varepsilon}$H]*m-i*, zHis[N$^{\delta}$H]*m-i* or the deprotonation of another hydrogen in the imidazole ring for zHis[N$^{\varepsilon}$HN$^{\delta}$H]*m-i*. To be more specific for zHis[N$^{\varepsilon}$HN$^{\delta}$H]*m-i*, pK_{a1} corresponds to the deprotonation at the N$^{\varepsilon}$ site. The results are given in Table 2. The pK_{a1} values for the canonical tautomers range from 17.4 to 19.3 and are similar to the theoretical results of Hudaky and Perczel [45]. The experimental estimate of the pK_a for histidine is 6.1 [46] and the experiments on histidine models that are methylated at the N$^{\varepsilon}$ or N$^{\delta}$ position show that the pK_a value is 6.61 for N$^{\delta}$ and 5.99 for N$^{\varepsilon}$ [24–27]. As shown in Table 2, among all the tautomers examined, the pK_a values for zHis[N$^{\varepsilon}$HN$^{\delta}$H]*m-i* are in best agreement with the experiments. This is not surprising since the pK_a measurement corresponds to the histidine structure with a protonated imidazole ring. As can be seen from Table 2, the agreement with the experiment improves with the increasing number of explicit water molecules. Like the case for the IR spectra discussed above, a full agreement with the experimental pK_a values is not expected due to the small number of explicit water molecules considered in the calculations. However, the pK_a values obtained for zHis[N$^{\varepsilon}$HN$^{\delta}$H] with two explicit water molecules are in satisfactory agreement with the experiments. It is interesting to note that our pK_a results for zHis[N$^{\varepsilon}$HN$^{\delta}$H] with two explicit water molecules are of comparable quality with the result of Ivaylo et al. [46] at pK_a = 6.8 that is found by a Car-Parrinello molecular dynamics simulation with a ring protonated histidine surrounded by 49 water molecules.

Table 2 pK_a values of histidine-$(H_2O)_m$ (m = 0,1,2)

Tautomers	m = 0 [Bare]		m = 1 [–H_2O]		m = 2 [–$(H_2O)_2$]	
	pK_{a1}	pK_{a2}	pK_{a1}	pK_{a2}	pK_{a1}	pK_{a2}
His[N$^{\varepsilon}$H]*m-i*	18.9	9.8	17.4	8.8	18.1	11.4
His[N$^{\delta}$H]*m-i*	18.6	8.7	18.7	11.8	19.3	10.1
zHis[N$^{\varepsilon}$H]*m-i*	17.4	11.5	17.9	13.8	21	13.2
zHis[N$^{\delta}$H]*m-i*	14.8	9.8	15.7	11.6	16.8	9.7
zHis[N$^{\varepsilon}$HN$^{\delta}$H]*m-i*	4.8	5.1	5.3	5.8	6.7	6

pK_{a1} corresponds to the deprotonation at the imidazole ring and pK_{a2} corresponds to the deprotonation at the other possible site

4 Summary

The possible microsolvation patterns of amino acids hydrated with one and two water molecules are depicted and the strategy of exploring the hydrated structures is described in detail. The hydrated structures of each tautomer of both canonical and zwitterionic histidines have been systematically explored and the effects of microsolvation and aqueous solvation are examined. The binding of the water molecules to the solute molecules may significantly alter their relative stability either in gas phase, in water clusters or in its solution phase. The lowest energy conformations of doubly hydrated histidine are more likely been found by trial structures generated by attaching two

water molecules simultaneously to the bare histidine rather than by the stepwise hydration method. The solvent effects combining the cluster model and the continuum model was also investigated. The results show that the zwitterions are more dominant in solution with more explicit water molecules attached. Attaching explicit water molecules is essential for the quantitative understanding of the solvent effects.

Hydrogen bond energies for the most stable structures of hydrated canonical and zwitterionic conformers of histidine have been analyzed by the AIM theory based upon the B3LYP/6-311++G(*d*,*p*) electron density. The red shift of N^{ε}-H stretching vibration of the imidazole ring for doubly hydrated structure of His[N^{ε}H] tautomer are observed with enhanced intensity and similar behavior has been detected for N^{δ}-H stretching in both singly and doubly hydrated form of His[N^{δ}H] tautomer. The OH stretching of COOH is also red shifted with enhanced intensity in their hydrated form.

The IR spectra of zwitterionic tautomers with two explicit water molecules are presented. Comparison with the experimental results at different pH solutions is discussed. Such comparison is shown to be useful to deduce the local hydration structures of histidine in solution.

The pK_a values for selected conformations of all tautomers of histidine with 0, 1, and 2 explicit water molecules are determined. The pK_a values obtained for the ring protonated zwitterions with two explicit water molecules are in good agreement with the experiments.

Acknowledgments Authors are thankful to the financial support of National Science Fund for Fostering Talents in Basic Science (J0630319) and the Knowledge Innovation Program of the Chinese Academy of Sciences (KJCX1.YW.07).

References

1. Jonsson PG, Kvick A (1972) Acta Crystallogr B 28:1827. doi:10.1107/S0567740872005096
2. Albrecht G, Corey RB (1939) J Am Chem Soc 61:1087. doi:10.1021/ja01874a028
3. Zwier TS (2001) J Phys Chem A 105:8827. doi:10.1021/jp011659+
4. Shishkin OV, Sukhanov OS, Gorb L, Leszczynski J (2002) Phys Chem Chem Phys 4:5359. doi:10.1039/b205351a
5. Robertson EG, Simons JP (2001) Phys Chem Chem Phys 3:1. doi:10.1039/b008225m
6. Sicinska D, Paneth P, Truhlar DG (2002) J Phys Chem B 106(10):2708. doi:10.1021/jp013252a
7. Schutz M, Burgi T, Leutwyler S, Fischer T (1993) J Chem Phys 98:3763. doi:10.1063/1.464055
8. Ebata T, Hashimoto T, Ito T, Inokuchi Y, Altunsu F, Brutschy B, Tarakeshwar P (2006) Phys Chem Chem Phys 8:4783
9. Brondsted NS, Andersen LH (2006) Biophys Chem 124:229. doi:10.1016/j.bpc.2006.04.002
10. Klamt A, Schüürmann G (1993) J Chem Soc Perkin Trans 2:799. doi:10.1039/p29930000799
11. Andzelm J, Kölmel C, Klamt A (1995) J Chem Phys 103:9312. doi:10.1063/1.469990
12. Barone V, Cossi M (1998) J Phys Chem A 102:1995. doi:10.1021/jp9716997
13. Cossi M, Rega N, Scalmani G, Barone V (2003) J Comput Chem 24:669. doi:10.1002/jcc.10189
14. Bandyopadhyay P, Gordon MS, Mennucci B, Tomasi J (2002) J Chem Phys 116:5023. doi:10.1063/1.1433503
15. Tortonda FR, Pascual-Ahuir JL, Silla E, Tunon I (1996) Chem Phys Lett 260:21. doi:10.1016/0009-2614(96)00839-1
16. Tortonda FR, Silla E, Tunon I, Rinaldi D, Ruiz-Lopez MF (2000) Theor Chem Acc 104:89. doi:10.1007/s002149900109
17. Tunon I, Silla E, Ruiz-Lopez MF (2000) Chem Phys Lett 321:433. doi:10.1016/S0009-2614(00)00365-1
18. Lee KT, Sung J, Lee KJ, Kim SK, Park YD (2002) J Chem Phys 116:8251. doi:10.1063/1.1477452
19. Lee K, Sung J, Lee K, Kim S, Park Y (2003) Chem Phys Lett 368:262. doi:10.1016/S0009-2614(02)01850-X
20. Jeon I-S, Ahn D-S, Park S-W, Lee S, Kim B (2005) J Quantum Chem 101:55 doi:10.1002/qua.20269
21. Park SW, Im S, Lee S, Desfrancois C (2007) Int J Quantum Chem 107: 1316. doi:10.1002/qua.21255
22. Ahn D, Park S, Jeon I, Lee M, Kim N, Han Y, Lee S (2003) J Phys Chem B 107:14109. doi:10.1021/jp031041v
23. Ellzy MW, Jensen JO, Hameka HF, Kay JG (2003) Spectrochim Acta A Mol Biomol Spectrosc 59:2619. doi:10.1016/S1386-1425(03)00036-2
24. Reynolds WF, Peat IR, Freedman MH, Lyerla JR (1973) J Am Chem Soc 95:328. doi:10.1021/ja00783a006
25. Boschcov P, Seidel W, Muradian J, Tominaga M, Paiva ACM, Juliano L (1983) Bioorg Chem 12:34. doi:10.1016/0045-2068(83)90005-6
26. Tanokura M (1983) Biochim Biophys Acta 742:576
27. Huang Z, Yu W, Lin Z (2006) J Mol Struct Theochem 801:7. doi:10.1016/j.theochem.2006.08.053
28. Huang Z, Lin Z, Song C (2007) J Phys Chem A 111:4340. doi:10.1021/jp067280a
29. Fei W, Rai AK, Lu Z, Lin Z (2009) J Mol Struct Theochem 895:65. doi:10.1016/j.theochem.2008.10.017
30. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas Ö, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Salvador P, Dannenberg JJ, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komáromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople J (1998) A Gaussian 98. Gaussian Inc, Pittsburgh
31. Scott AP, Radom L (1996) J Phys Chem 100:16502. doi:10.1021/jp960976r
32. Chen M, Lin Z (2007) J Chem Phys 127(15):154314. doi:10.1063/1.2777161
33. Rappé AK, Casewit CJ, Colwell KS, Goddard WAIII, Skiff WM (1992) J Am Chem Soc 114:10024. doi:10.1021/ja00051a040
34. Bader RFM (1990) Atoms in molecules: a quantum theory. Clarendon Press, Oxford
35. Popelier PLA (2000) Atoms in molecules: an introduction. Prentice-Hall, Harlow
36. Liptak MD, Gross KC, Seybold PG, Steven F, Shields GC (2002) J Am Chem Soc 124:6421. doi:10.1021/ja012474j

37. Leng Y, Zhang M, Song C, Chen M, Lin Z (2008) J Mol Struct Theochem 858:52. doi:10.1016/j.theochem.2008.02.016
38. Aiken M, Gordon MS (2006) J Am Chem Soc 128:12835. doi:10.1021/ja062842p
39. Bonaccorsi R, Palla P, Tomasi J (1984) J Am Chem Soc 106:1945. doi:10.1021/ja00319a008
40. Ding Y, Krogh-Jespersen K (1992) Chem Phys Lett 199:261. doi:10.1016/0009-2614(92)80116-S
41. Espinosa E, Molins E, Lecomete C (1998) Chem Phys Lett 285:170. doi:10.1016/S0009-2614(98)00036-0
42. Mesu JG, Visser T, Soulimani F, Weckhuysen BM (2005) Vib Spectrosc 39:114. doi:10.1016/j.vibspec.2005.01.003
43. Deplazes E, van Bronswijk W, Zhu F, Barron LD, Ma S, Nafie LA, Jalkanen KJ (2008) Theor Chem Acc 119:155. doi:10.1007/s00214-007-0276-8
44. Ling S, Yu W, Huang Z, Lin Z, Haranczyk M, Gutowski M (2006) J Phys Chem A 110:12282. doi:10.1021/jp0645115
45. Peter H, Andras P (2004) J Phys Chem A 108:6195. doi:10.1021/jp048964q
46. Ivaylo I, Bin C, Simone R, Klein ML (2006) J Phys Chem B 110:6365. doi:10.1021/jp056750i