

Vassil B. Delchev · Hans Mikosch

# Theoretical study of the intermolecular H-bonding and intermolecular proton transfer between isocytosine tautomeric forms and *R,S*-lactic acid

Received: 17 August 2005 / Accepted: 9 November 2005 / Published online: 19 May 2006  
© Springer-Verlag 2006

**Abstract** Eight H-bonded complexes between isocytosine (*isoC*) tautomeric forms and *R/S*-lactic acid (LA) have been studied at the B3LYP and HF levels of theory using 6–31+G(d) basis set. The energy barriers of the intermolecular proton transfers were also estimated as the results showed that they are several times lower than those of the intramolecular proton transfers of *isoC* in the gas phase. Furthermore, the energy barriers of the tautomerizations in which the carboxylic H-atom takes part are several times lower than those in which the LA OH group assists the proton transfer.

**Keywords** Density functional methods · H-bonding · Isocytosine · Proton transfer

## Introduction

One specific role of the DNA bases is connected with their ability to establish noncovalent H-bonds that are responsible for the encoding and the expression of genetic information. Guanine is one of the most studied DNA bases by means of experimental [1–4] and theoretical [5–9] techniques. It exists in several tautomeric forms [10, 11], which are in dynamic equilibrium. One very important fragment from the guanine skeleton is the pyrimidine residue, which corresponds to isocytosine (*isoC*). Isocytosine has not been detected in natural DNA and RNA, but its C-nucleoside has been used successfully as an antileukemia drug [12]. Five possible tautomers of *isoC* have been suggested in the work of Gorb et al. [13]. However, only

two of them have been determined as most stable in the gas phase [13]: the so-called amino oxo- and amino hydroxo-*isoC*. In several theoretical works of Gorb et al. [13–16] it has been shown that the interaction of the DNA bases with a water molecule is able to change the stability order of their tautomers (as compared to the isolated molecules). We expect the same trend to be observed when nucleobases and isocytosine are bound by intermolecular H-bonds with other small protic organic molecules, such as for example lactic acid (LA).

Using the HF/6–31G(d,p) method, Shishkin [17] has established that *isoC* has a planar structure. However, several previously published investigations have revealed that the amino groups in cytosine, isocytosine and guanine can adopt a non-planar or so-called pyramidal geometry [18–21].

There are two principal ways for transformation of the amino oxo-tautomer into the amino hydroxo-one: intramolecular and intermolecular proton transfer, the latter assisted by protic molecules. It is well known that the first type of proton transfer passes through high energy barriers [6, 7, 13, 22, 23], over 200 kJ mol<sup>-1</sup>, whereas the second one occurs with a several times lower barrier [7]. So far, only water molecules have been taken into account to assist the intermolecular hydrogen transfer in isocytosine [13]. However, it is interesting to know how other protic molecules, abundant in the living world, support these proton transformations. Such a molecule is, for example, lactic acid. The interest in this molecule follows from the fact that this compound has two enantiomers: *R*- and *S*-lactic acid.

The aim of this paper is:

- 1) To perform a theoretical study at the B3LYP and HF levels on the possible H-bonded systems between *R/S*-lactic acid and *isoC* in positions appropriate for oxo-hydroxo tautomerism;
- 2) to estimate the difference in the energy barriers of the intermolecular proton transfers assisted by *R*-LA and *S*-LA. As a model, we chose the amino oxo- and amino hydroxo-tautomers of *isoC*, which are an explicit example for oxo-hydroxo tautomerization.

V. B. Delchev (✉)  
Department Physical Chemistry, University of Plovdiv,  
Plovdiv, Bulgaria  
e-mail: vdelchev@pu.acad.bg

H. Mikosch  
Institute of Chemical Technologies and Analytics,  
Technical University Vienna,  
Getreidemarkt 9,  
A-1060 Wien, Austria

The DFT hybrid functional B3LYP along with the 6–31+G(d) basis set has shown good accuracy for studying such processes [24–27] with comparatively low cost in computation time [28].

### Computational details

The calculations were performed at the restricted Hartree–Fock and DFT levels using the 6–31+G(d) basis set and the program Gaussian 98 [29]. Both methods were used to optimize the structures of the supersystems and monomers. Frequency calculations were performed to prove that the resulting stationary points correspond to minima. The mechanisms of the proton transfers were studied using the geometries of two minima and the QST2 procedure implemented within GAUSSIAN 98 [29].

The standard free energies, enthalpies, entropy factors, and equilibrium constants of the complexations were found using the thermodynamic equations and  $\Delta G = \Delta H - T\Delta S$  and  $K = e^{-\Delta G/R}$  [30, 31]. The rate constants of the forward and reverse reactions of the proton transfers were calculated by the Eyring equation [31] within the adiabatic approximation (with the transmission coefficient equal to unity).

The bonding energy  $\Delta E_b$  and BSSE for each supersystem (complex) were estimated according to the equations [32–37]

$$\begin{aligned}\Delta E_b &= E_{SS} - (E'_{isoC} + E'_{LA}) \text{ and } \Delta BSSE \\ &= \sum_i^n (E_{m_i} - E'_{m_i}),\end{aligned}$$

with  $E_{SS}$  as the energy of the supersystem;  $E'_{isoC}$  and  $E'_{LA}$  are energies of the monomers (*isoC* and *LA*) calculated with so-called ‘ghost’ orbitals of the other monomer, and  $E_{m_i}$  are the energies of the individual monomers frozen in their aggregate geometries (single-point).

The energy of the *isoC*-*LA* interaction ( $\Delta E_{int}$ ) was calculated as the difference between the energies of the complex and isolated *isoC* and *LA* molecules. The difference between the energies of the complex and the deformed configurations of *isoC* and *LA* in the complex ( $\Delta E$ ) was found using the single-point energies of the monomers in their aggregate configurations.

## Results and discussions

### Geometry of the structures in the local minima

The theoretical investigation was done with conformational analysis of *R*- and *S*-*LA* looking for the most stable conformer in a complexation with *isoC*. The conformational flexibility of *R*- and *S*-*LA* was checked by scan of the dihedral angles  $\angle H(1)O(2)C(3)C(9)$  and  $\angle H(12)O(11)C(9)C(3)$  with a scan range of  $\pm 180^\circ$  and step of  $10^\circ$ . With

respect to the  $C(\alpha)$ – $C(\beta)$  bond the most stable should be the *gauche* conformation (Scheme 1).

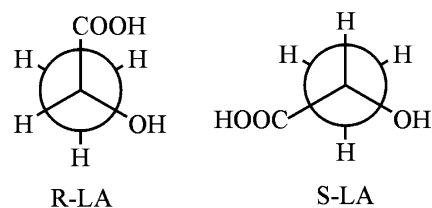
Both methods used showed that the H atom from the  $C(\alpha)$ –OH group forms a weak intramolecular H-bond with the carbonyl O atom from the carboxylic function. These most stable conformations are shown in Fig. 1 (B3LYP/6–31+G(d)).

The calculations at the B3LYP level predicted that the rotation of the OH group in *R*-*LA* around  $O(2)C(3)$  has an asymmetric curve with no other stable conformers corresponding to minima. There is only one unstable form (R1) in whose structure H(1) is most distant from O(11). This conformation is not stable (no energy minimum) because of structural hindrance between the H(1) atom and the  $CH_3$  group. The full rotation (at  $360^\circ$ ) of the  $H(1)O(2)$  group occurs with an energy barrier of about  $21 \text{ kJ mol}^{-1}$  ( $22 \text{ kJ mol}^{-1}$  predicted by HF calculations). The same rotation in *S*-*LA* led to one more minimum corresponding to a structure (S1), in which the atoms O(11), C(9), C(3) and O(2) are placed almost in one plane:  $\angle O(11)C(9)C(3)O(2) = 176.1^\circ$ . In the R1 conformer  $\angle O(11)C(9)C(3)O(2) = 26.3^\circ$ , as the repulsion between the electron pairs of the O(11) and O(2) is large. The full rotation of the  $H(1)O(2)$  in *S*-*LA* requires an energy barrier of  $27 \text{ kJ mol}^{-1}$  ( $28 \text{ kJ mol}^{-1}$  predicted by HF calculations). Structures R1 and S1 are about 19 and  $25 \text{ kJ mol}^{-1}$ , respectively, less stable than the conformers depicted in Fig. 1.

The full rotation of the carboxylic OH group requires up to three times more energy to occur:  $68 \text{ (R-LA)}$  and  $72 \text{ kJ mol}^{-1}$  (*S*-*LA*). The energy barriers obtained by HF method are 68 and  $74 \text{ kJ mol}^{-1}$ , respectively. Each curve has three minima. The deepest minimum corresponds to the most stable conformer (shown in Fig. 1). The second minimum is an intermediate structure denoted as R2 (S2 for *S*-*LA*). Its energy is  $54 \text{ kJ mol}^{-1}$  ( $56 \text{ kJ mol}^{-1}$  for S2) higher than that of the most stable conformer. The third corresponds to conformer R3 (or S3) having a  $62 \text{ (}72 \text{ kJ mol}^{-1}$  for S3) higher energy than the most stable conformers.

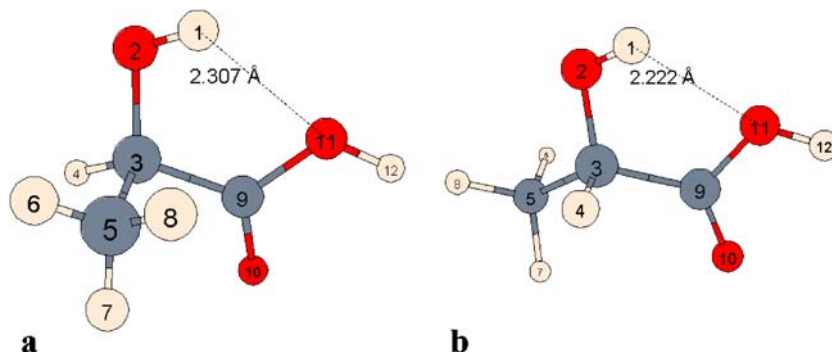
The dihedral angles  $H(12)O(11)C(9)C(3)$  in the most stable *R*- and *S*-*LA* conformers (Fig. 1) are  $177.6$  and  $176.9^\circ$ , respectively. These angles are  $43.4$  and  $53.0^\circ$  in the R2 and S2 conformers, and  $26.3$  and  $37.0^\circ$  in the structures R3 and S3.

All initial structures for geometry optimizations of the hydrogen-bonded complexes studied here were obtained by circling one isocytosine tautomer around the most stable *R*- or *S*-lactic acid conformer (Fig. 2). As a result eight, possible H-bonded supersystems are obtained. They were



**Scheme 1** Staggered conformations of *R*- and *S*-*LA*

**Fig. 1** Optimized conformers of **aR**-, and **bS**-LA (B3LYP/6-31+G(d)) in the gas phase

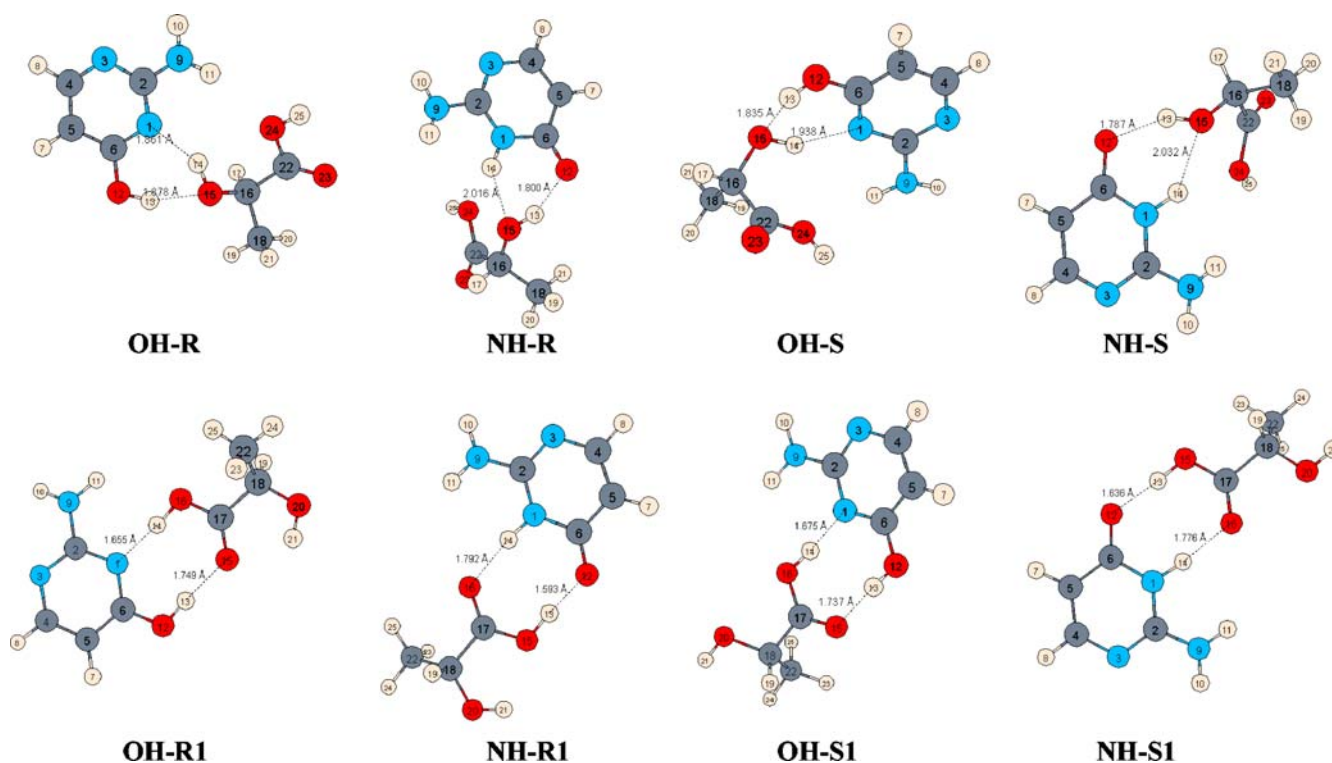


fully optimized at both theoretical levels. The B3LYP geometries are shown in Fig. 2.

Following the H-bond-classification given in [38], the two H-bonds in the OH-R and OH-R1 supersystems are “medium” (Y...H 1.6–1.9 Å). However the H-bonds in the OH-R1 supersystem are 0.206 and 0.129 Å shorter, which can be explained by the greater affinity of the carboxylic function to form intermolecular H-bridges. The Hartree-Fock method predicted quite longer H-bonds,  $O_{15}\dots H_{13}=2.006$  Å and  $N_1\dots H_{14}=2.062$  Å for the OH-R supersystem and  $O_{15}\dots H_{13}=1.877$  Å and  $N_1\dots H_{14}=1.875$  Å for OH-R1 one. Obviously the reason is that HF method does not take electronic correlation into account, whereas the DFT method more accurately renders an account of the electron density distribution. A detailed discussion of the pyramidal character of the amino group in the nucleobases is given in the work of Chandra, Delchev, van Mourik et al. [39–41].

At first sight, the NH-R complex should be less stable than the OH-one since one of the H-bonds is longer than 1.9 Å. However, the stability of the complex comes from the next H-bond, i.e.  $H_{13}\dots O_{12}$ , which is unusually short for this kind of H-bonding. The same H-bond in the NH-R1 supersystem is 1.593, 0.207 Å longer. This bond is strong ( $Y\dots H < 1.6$  Å) [38] and thus determines the high stability of this supersystem.

One of the bonds in the OH-S (Fig. 2) is weak, whereas the other is medium. Shorter intermolecular H-bonds were calculated in the OH-S1 supersystem:  $H_{14}\dots N_1=1.675$  Å and  $H_{13}\dots O_{15}=1.737$  Å. Unfortunately, the combination HF/6-31+G(d) was not able to find a local minimum for the OH-S supersystem, even when very tight convergence criteria were applied. The predicted structure gives one imaginary vibration with a negative vibrational eigenvalue at  $-9$   $\text{cm}^{-1}$  and very low intensity ( $0.68$   $\text{km mol}^{-1}$ ).



**Fig. 2** Optimized structures of the hydrogen-bonded supersystems composed between isocytosine tautomers and *R,S*-lactic acid (B3LYP)

One of the H-bonds in the NH-S is very short, whereas the other is longer than 2 Å. The two intermolecular H-bonds in the NH-S1 supersystem are 0.151 and 0.256 Å shorter than in the NH-S one. As expected, the intermolecular H-bonds in the supersystems formed between the carboxylic function of the *R/L*-LA and the *isoC* monomer are stronger than those formed between the OH group of *R/S*-LA and *isoC*.

In all supersystems, the carboxylic OH group lies in the vicinity of the NH bonds of the amino group of *isoC* because comparatively weak intermolecular H-bonds (~2.6 Å) are formed between O(24) and H(14) in the NH forms and between O(24) and H(11) from the amino group in the OH forms. These bonds are strong enough to overcome the repulsion between the carboxylic OH and NH bonds. It should be mentioned that in the supersystems the LA conformation is changed relative to its most stable conformations discussed above for free *R,S*-LA. In the structures of the complexes the LA conformation is not typical for a free molecule in the gas phase, but it is the most stable one according to the intermolecular H-bonds that it forms.

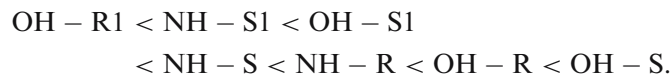
### Thermodynamic parameters of complex formations

Gorb et al. [13] have discussed the finding that the HF calculations predicted a lower energy for the amino hydroxo-form than for the amino oxo-form of *isoC*. However, the more accurate B3LYP method gave the reversed stability order for these *isoC* tautomers.

The relative energies and relative zero-point energies of the complexes studied together with some thermodynamic parameters of the complexation processes are listed in Table 1.

The data from Table 1 clearly show that the supersystem NH-R1 should be the most stable, not only because it is formed by the shortest H-bond between the monomers. We think that the high stability of this form may result from the

high stability of the NH monomer of *isoC* (this will be discussed below). The remaining supersystems follow this stability pattern:



The least stable supersystem is composed between the OH monomer of *isoC* and *S*-lactic acid. The complexation process of OH-S would have the lowest equilibrium constant (largest, Table 1), which shows that the equilibrium between the monomers and the supersystem, in the gas phase, is shifted toward the monomers, despite the comparatively high steric (entropy) factor for this complexation process.

Both methods showed that all complex formations are enthalpically favored (exothermic). They are accompanied by a considerable heat effect. Similarly to the guanine-water and guanine-methanol complexes described in [5, 39], the entropy factor  $T\Delta S_{298}^{\circ}$  of each complexation is negative. For the supersystems formed between the OH group of *R/S*-LA, it is clear that the absolute value of the entropy factor is higher than the absolute value of the enthalpy. Therefore, the steric effects dominate during the complexations. However, the complexations between carboxylic group and *isoC* are accompanied by larger energy (enthalpy) changes than steric (entropy) alterations. These complexations have a negative variation of the free Gibbs energy, which means that they are thermodynamically favored, and the equilibrium is shifted toward the supersystem. On the other hand, the values of the complexations between the OH group of *R/L*-LA and *isoC* are positive. However, the values are in the uncertainty interval -42 to +42 kJ mol<sup>-1</sup>. Therefore, it remains questionable if the complexations are thermodynamically favored or not. These findings are in agreement with the thermodynamic parameters of the supersystems *isoC*-water [13].

**Table 1** Relative energies ( $\Delta E$ ) and relative zero-point energies ( $\Delta E^0$ ) of the supersystems, and thermodynamic parameters of the complex formations (kJ mol<sup>-1</sup>)

	$\Delta E$		$\Delta E^0$		$\Delta G_{298}^{\circ}$		$\Delta H_{298}^{\circ}$		$T\Delta S_{298}^{\circ}$	
	1	2	1	2	1	2	1	2	1	2
OH-R	44	40	46	40	9	17	-35	-24	-43	-43
NH-R	37	33	39	34	2	7	-41	-33	-43	-39
OH-S	54	51	56	50	19	35	-24	-16	-43	-50
NH-S	36	31	37	32	0	6	-42	-34	-42	-40
OH-R1	11	12	10	12	-24	-9	-70	-52	-46	-43
NH-R1	0 <sup>a</sup>	0 <sup>b</sup>	0 <sup>c</sup>	0 <sup>d</sup>	-33	-23	-80	-67	-47	-44
OH-S1	30	30	29	30	-6	9	-51	-34	-45	-42
NH-S1	13	12	12	11	-21	-11	-67	-55	-46	-44

1: B3LYP/6-31+G(d); 2: HF/6-31+G(d)

For each column the energy

<sup>a</sup>-738.605604;

<sup>b</sup>-734.363064;

<sup>c</sup>-738.410255;

<sup>d</sup>-734.151610 a.u.

was taken as a conventional zero

As concerns the two methods used in this work, one can say that DFT predicted lower absolute energies and thermodynamic parameters of the supersystems compared to HF. The average difference between predicted energies from DFT and HF is  $11133 \text{ kJ mol}^{-1}$ .

In order to understand the interacting forces that hold together the two monomers in one supersystem better, we calculated the hydrogen-bonding energies and bonding parameters (Table 2).

As one can see from Table 2, the NH–R1 supersystem has the highest bonding energy, followed by the NH–S1 one. As a whole, the bonding energies and the interaction energies of the supersystems are larger for the complexes constituted by H-bonding between the carboxylic function of the *R/S*-LA and *isoC*. Among them the complexes of the NH-type are the most stable systems. Obviously, the complexations with the OH-form of *isoC* lead to the supersystems with lower stability.

A clear trend is observed (Table 2): lower stability (bonding and interaction energy)–higher  $\Delta(\text{BSSEs})$  and vice versa. The B3LYP calculations showed (Table 2) that the most stable *isoC* monomer frozen in its aggregate geometry (SP) is found within the NH–S supersystem. According to the HF method, this monomer is available

within the OH–R supersystem. This fact clearly shows that the high stability of the NH–R1 supersystem does not come from the high stability of its constituting monomer.

The large difference between the  $\Delta E_{\text{int}}$  and  $\Delta E$  shows that the monomers within the supersystems are with rather changed configurations (especially LA), far from their equilibrium geometries (isolated monomers).

Comparing the bonding energies (Table 2) of the supersystems and the conformational energies of *R*- and *S*-LA, it is seen that the LA molecules can change their conformation very easily (with respect to H(1)O(2) and H(12)O(11)) to create the most favorable conditions for intermolecular interaction. The structural analysis of the supersystems showed that the OH functions are more flexible with respect to creating of intermolecular H-bonds. However, the bonding energies are not enough to reach conformers R2, R3 and S2, S3, whereas they are adequate for reaching the conformers R1 and S1.

A very good correlation (B3LYP) was found between the intermolecular H-bonds H(14)...N(1) and the bonding energies of the OH-supersystems:  $\Delta E_b = 153.73(\text{H}(14)\dots\text{N}(1)) - 330.22$ ;  $r^2 = 0.9778$ . The correlation between  $\Delta E_b$  and H(13)...O(15) has a slightly lower correlation coefficient: 0.8806.

**Table 2** *isoC*–LA interaction energies and their components ( $\text{kJ mol}^{-1}$ )

	$E_{m_i}$				E'			
	<i>isoC</i>		LA		<i>isoC</i>		LA	
	1	2	1	2	1	2	1	2
OH–R	0.62	0.00 <sup>b</sup>	1.63	5.20	0.00 <sup>c</sup>	0.41	1.61	5.41
NH–R	0.02	2.46	6.96	13.76	0.92	3.56	7.41	14.39
OH–S	1.19	0.45	6.89	10.70	0.39	0.00 <sup>f</sup>	6.80	10.82
NH–S	0.00 <sup>a</sup>	2.40	7.00	12.34	0.99	3.58	7.33	12.79
OH–R1	4.18	2.41	0.00 <sup>c</sup>	0.00 <sup>d</sup>	3.48	1.96	0.00 <sup>g</sup>	0.00 <sup>h</sup>
NH–R1	3.44	4.56	10.23	9.50	4.07	5.27	10.77	9.99
OH–S1	4.22	2.35	18.56	19.22	3.66	5.31	18.76	19.19
NH–S1	3.18	4.65	20.83	23.05	3.77	2.04	21.34	23.29
	$\Delta E_b$		$\Delta E_{\text{int}}$		$\Delta E$		$\Delta(\text{BSSE})$	
	1	2	1	2	1	2	1	2
OH–R	–39.99	–26.76	–41.55	–31.88	–46.49	–33.31	6.49	6.55
NH–R	–53.67	–45.19	–47.72	–40.19	–58.18	–50.63	4.51	5.44
OH–S	–35.16	–20.91	–30.97	–20.52	–41.90	–28.42	6.74	7.50
NH–S	–55.15	–45.64	–49.04	–41.71	–59.68	–51.18	4.53	5.54
OH–R1	–74.86	–69.50	–74.55	–59.21	81.42	–57.85	6.56	7.63
NH–R1	–97.21	–75.94	–84.74	73.63	–101.89	–81.90	4.68	5.97
OH–S1	–74.96	–51.52	–55.54	–40.71	–81.18	–59.03	6.22	7.51
NH–S1	–94.98	–77.26	–72.07	61.11	–99.73	–83.54	4.76	6.27

1: B3LYP/6–31+G(d); 2: HF/6–31+G(d)

The absolute energy

<sup>a</sup>–394.949587;

<sup>b</sup>–392.630699;

<sup>c</sup>–343.622408;

<sup>d</sup>–341.706518;

<sup>e</sup>–394.950455;

<sup>f</sup>–392.631826;

<sup>g</sup>–343.623770;

<sup>h</sup>–341.708124 a.u.

was taken as a conventional zero

For the NH-supersystems, very good correlations exist between  $\Delta E_b$  and H(13)...O(12) and H(14)...O(15/16):  $\Delta E_b=227.51(\text{H}(13)\dots\text{O}(12))-462.93$ ;  $r^2=0.9825$ ;  $\Delta E_b=172.41(\text{H}(14)\dots\text{O}(15/16))-403.51$ ;  $r^2=0.9876$ .

Without the diffuse functions in the basis set, the energies of the supersystems are increased by about 97 kJ mol<sup>-1</sup> according to the B3LYP optimizations and 57 kJ mol<sup>-1</sup> according to the HF method. The use of diffuse functions for the calculations is desirable since the basis set superposition errors (BSSE) are known to decrease with basis set size and are reduced by the inclusion of diffuse functions [42, 43]. The diffuse functions allow orbitals to occupy a larger region of space. Therefore, basis sets with diffuse functions are important for systems where electrons

are relatively far from the atomic nucleus [44], as in our hydrogen-bonded systems with many lone electron pairs.

#### Vibrational analysis of the supersystems

The calculations at the B3LYP level gave the asymmetric NH<sub>2</sub> stretching vibrations in the NH-and OH-*iso*C to be at 3678 (3840 cm<sup>-1</sup> HF) and 3721 (3967 cm<sup>-1</sup> HF) cm<sup>-1</sup>. The symmetric vibrations are at 3572 (3809 cm<sup>-1</sup> HF) and 3598 (3839 cm<sup>-1</sup> HF) cm<sup>-1</sup>. The characteristic C=O vibration of the NH-*iso*C tautomer is calculated at 1772 (1946 cm<sup>-1</sup> HF) cm<sup>-1</sup> (the same vibration measured experimentally between KBr discs is at 1620 cm<sup>-1</sup> [45]),

**Table 3** Harmonic vibrational frequencies (cm<sup>-1</sup>) and their intensities (km mol<sup>-1</sup>) for the supersystems in the range over 1500 cm<sup>-1</sup> (by the B3LYP/6-31+G(d) resulted values)

OH-R		OH-S		NH-R		NH-S		Exp. [45]	Ass
1	2	1	2	1	2	1	2		
3716/93	3962/80	3716/63	3965/92	3691/51	3932/98	3694/53	3934/325	–	$\nu_{as}NH_2$
3589/106	3834/85	3594/63	3837/69	3578/68	3815/83	3580/70	3816/83	–	$\nu_sNH_2$
3445/1358	3974/955	3453/1413	3992/512	3420/838	3925/464	3424/890	3932/328	3261	$\nu_{as}OH$
3343/338	3931/104	3388/48	3930/354	3357/286	3767/207	3361/380	3767/205	2990	$\nu_sOH$
3154/13	3314/19	3144/26	3293/17	3151/12	3314/16	3148/1	3304/20	–	$\nu_{as}CH_3^{LA}$
3142/16	3298/26	3127/26	3278/21	3131/20	3285/35	3133/22	3290/38	–	$\nu_{as}CH_3^{LA}$
3071/14	3230/18	3055/17	3213/18	3061/21	3216/14	3061/19	3218/20	2943	$\nu_sCH_3^{LA}$
2978/28	3177/29	3118/1	3321/31	3056/18	3244/21	3052/19	3246/16	2833	$\nu CH^{LA}$
1828/323	2021/418	1833/320	2028/431	1812/313	2000/414	1815/315	2005/423	1732	$\nu C=O^{LA}$
–	–	–	–	1742/773	1908/1123	1740/771	1907/1106	1620	$\nu C=O^{isoC}$
1674/500	1820/722	1670/473	1818/718	1681/478	1834/579	1681/479	1833/576	–	$\delta NH_2, \rho H_{13}$
1656/65	1803/76	1652/69	1799/79	1635/119	1798/104	1634/115	1797/103	–	$\delta NH_2, \rho H_{13}$
1529/382	1625/357	1524/192	1618/337	–	–	–	–	–	$\nu C-NH_2, \delta_i$
1518/15	1634/19	1516/12	1635/26	1521/8	1637/7	1520/6	1638/2	–	$\delta CH_3^{LA}$
1509/7	1628/19	1510/6	1631/44	1519/44	1635/25	1510/10	1629/10	–	$\delta CH_3^{LA}$
1505/221	1649/187	1503/286	1651/210	1471/267	1580/253	1469/145	1580/47	–	$\nu C-OH^{isoC}, \rho H_{13}$
150/1	397/29	142/6	386/32	177/7	396/44	165/12	394/43	–	$\nu_i$
OH-R1		OH-S1		NH-R1		NH-S1		Exp. [45]	Ass
1	2	1	2	1	2	1	2		
3722/73	3957/83	3717/72	3954/86	3711/75	3953/97	3710/77	3953/101	–	$\nu_{as}NH_2$
3671/79	4065/103	3751/23	4109/62	3734/56	4101/88	3755/26	4112/60	2990	$\nu OH^{LA}$
3601/82	3834/73	3597/83	3833/68	3591/93	3827/116	3589/96	3826/125	–	$\nu_sNH_2$
3315/1676	3876/1094	3267/1814	3847/1211	2799/2369	3616/1005	2933/1898	3656/102	3261	$\nu O-H_{13}$
3187/25	3380/107	3185/31	3378/27	3190/24	3386/22	3188/27	3384/23	–	$\nu C-H_8$
3144/15	3302/25	3140/13	3298/23	3152/17	3312/22	3142/12	3304/19	–	$\nu_{as}CH_3^{LA}$
3135/21	3288/35	3117/24	3275/35	3138/21	3296/31	3117/23	3273/37	–	$\nu_{as}CH_3^{LA}$
3061/16	3217/18	3049/22	3209/27	3067/15	3228/20	3050/22	3209/34	2943	$\nu_sCH_3^{LA}$
3013/24	3241/16	3038/24	3241/20	2966/36	3166/38	3023/29	3217/23	2833	$\nu CH^{LA}$
2625/2769	3602/1391	2730/2441	3640/1236	3244/1344	3690/1144	3238/154	3691/1935	–	$\nu O/N-H_{14}$
1736/403	1947/513	1744/382	1949/499	1749/988	1945/830	1767/812	1970/798	1732	$\nu C=O^{LA}$
1679/481	1825/547	1679/470	1824/552	1685/574	1830/692	1685/567	1830/696	–	$\delta NH_2$
1653/215	1801/244	1652/209	1800/241	1644/97	1770/139	1644/99	1796/135	–	$\delta NH_2$
1625/240	1785/470	1626/280	1786/499	1611/42	1795/123	1612/40	1770/127	–	$\nu C_4-C_5$
1584/12	1664/47	1578/19	1662/85	1714/281	1889/846	1722/495	1892/951	1620	$\rho H_{14}, \nu C=O^{isoC}$
1520/33	1635/14	1525/63	1635/31	1522/45	1637/1	1528/55	1650/24	–	$\delta CH_3^{LA}$
153/4	147/1	148/1	140/2	160/12	145/7	159/9	144/6	–	$\nu_i$

1: B3LYP/6-31+G(d); 2: HF/6-31+G(d)

whereas the OH stretching of the OH-*iso*C tautomer at 3689 (4061  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ .  $\text{NH}_2$ -scissoring vibrations of the NH- and OH-*iso*C tautomers were calculated at 1681 (1836  $\text{cm}^{-1}$  HF) and 1669 (1820  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ , respectively.

As expected, both methods predicted similar vibration spectra of the lactic acid enantiomers. The H(1)O(2) stretching vibration is calculated at 3749 (4061  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ , whereas the carboxyl OH stretching vibration in the two enantiomers is at 3689 (4043  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ . Two bands of the asymmetric stretchings of the  $\text{CH}_3$  group were found at 3157 and 3142 (3304 and 3291  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ . The same symmetric vibrations are at 3071 (3220  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ .  $\text{C}(\alpha)\text{H}$  stretchings were calculated at 2969 (3210  $\text{cm}^{-1}$  HF)  $\text{cm}^{-1}$ . The characteristic  $\nu(\text{C}=\text{O})$  was found at 1827  $\text{cm}^{-1}$  (1997  $\text{cm}^{-1}$  HF), and measured at 1732  $\text{cm}^{-1}$  in a liquid film [45].

Several characteristic vibrations in the range over 1500  $\text{cm}^{-1}$  together with a few intermolecular stretching vibrations ( $\nu_i$ ) are listed in Table 3.

Data from Table 3 (the values are unscaled) clearly show that the DFT method predicts better values of the harmonic frequencies, closer to the experimental ones. For example, if the  $\nu_{\text{as}}\text{NH}_2$  stretching in the OH-R/S supersystems is

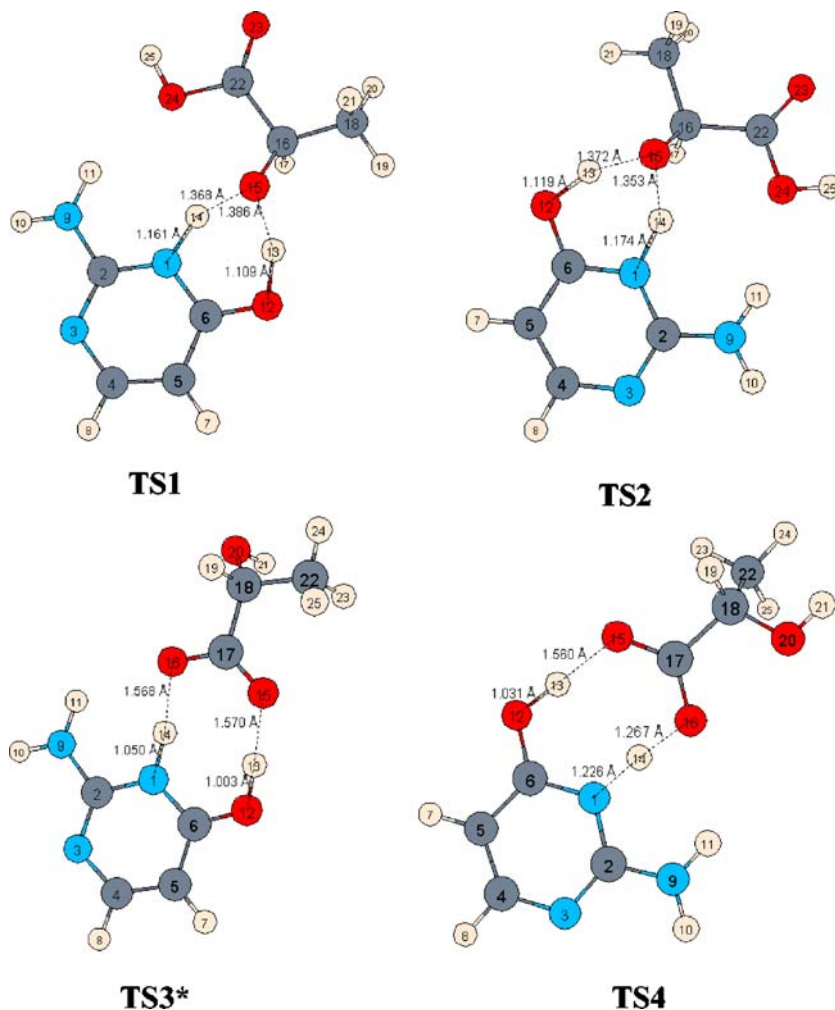
corrected with a scale factor of 0.9613 [44] it reproduces very well the experimental value with insignificant difference—only 5  $\text{cm}^{-1}$ . The reason for this small discrepancy between theory and experiment is that the experimental frequencies [2] were measured for free guanine in the gas phase. A larger difference between experimental and corrected theoretical (scaling factor 0.8929 [44]) frequency of the  $\nu_{\text{as}}\text{NH}_2$  stretching in the OH-R/S predicts HF method: about 38  $\text{cm}^{-1}$ .

The OH-vibrations are observed within the interval 3300–3800  $\text{cm}^{-1}$  (B3LYP) in the theoretical spectra of all supersystems. The most characteristic intermolecular vibrations were calculated in the low-frequency interval 130–160  $\text{cm}^{-1}$  corresponding to the spatial motion of the monomers along the intermolecular H-bonds.

The H-bonding in the NH-R/S supersystems causes a shift of the characteristic  $\nu\text{C}=\text{O}^{\text{isoC}}$  vibration (about 30 and 40  $\text{cm}^{-1}$  according to B3LYP and HF calculations, respectively) towards smaller wave numbers. Conversely,  $\nu\text{C}=\text{O}^{\text{LA}}$  is not affected by the H-bonding and the bands do not change their positions drastically in the monomers and supersystems.

Very large intensities (over 1000  $\text{km mol}^{-1}$ ) are calculated for vibrations along the intermolecular H-bonds

**Fig. 3** Optimized transition state structures of the of the intermolecular proton transfers (B3LYP). \*The transition state resulted by the HF calculations



( $\nu\text{O-H}_{13}$  and  $\nu\text{O/N-H}_{14}$ ) in the supersystems OH/NH-R1/S1. However, in their vibrational spectra, lower frequencies of the intermolecular motions were measured compared to these modes in the supersystems OH/NH-R/S. Thus, the intermolecular H-exchange between carboxylic group and *isoC* should have a lower energy barrier than the H-exchange including the OH group of *R/S-LA*.

### Intermolecular proton transfers

Knowledge of the inter- and intramolecular proton transfer mechanisms outlines the picture of the prototropic processes in the nucleic acid bases and their derivatives. It is well known that the intramolecular proton transfers in the DNA bases (and *isoC*) in the gas phase have quite high energy barriers [6, 7, 13]. These barriers undergo lowering if other protic molecules, e.g. water [7, 13], methanol [6] assist the proton transfer. The mechanism of the tautomeric equilibria in *isoC* assisted by water molecule has been studied in detail [13].

The structures of the transition states corresponding to intermolecular proton transfer assisted by *R,S*-lactic acid, found at the B3LYP/6-31+D(d) level, are shown in Fig. 3.

The transition state TS1 lies along the reaction coordinate connecting the supersystems OH-R and NH-R (OH-RNH-R). The structure TS2 connects the clusters OH-S

and NH-S (OH-SNH-S). The two transition states are first-order saddle points since in the vibrational spectrum of each, one parallel mode was calculated as an imaginary frequency. These parallel modes describe the intermolecular H-exchange.

Unfortunately, many attempts to optimize the transition state TS3 at the B3LYP level failed. This is why we must be satisfied with the transition state given by HF calculations. The transition states TS3\* (HF) and TS4 were found as first order saddle points on the potential hypersurfaces of the transformations OH-R1NH-R1 and OH-S1NH-S1.

The calculated energy barriers of the tautomeric conversions of *isoC* assisted by LA are given in Table 4.

Because it is more flexible, the carboxylic H atom is exchanged easier than the OH hydrogen atom of *R/S-LA*. The rate constants in the first case are larger than in the second. This reflects the energy barriers of the reactions, which are several times lower in the second case. Moreover, the transformations OH-R1/S1NH-R1/S1 have negative entropy factors and larger exothermicity. Comparing the rate constants and the energy barriers, one can see that the transformation OH-S1NH-S1 is the most favorable. In other words, *S-LA* facilitates to a greater extent the proton exchange processes in *isoC*.

As a whole, the energy barriers estimated here are several times (about four) lower than those found for the

**Table 4** Energies of the transition states ( $E_{\text{TS}}$  in a.u.), energy barriers, thermodynamic parameters, and kinetic data of the intermolecular proton transfers ( $\text{kJ mol}^{-1}$ )

	OH-R $\rightleftharpoons$ NH-R		OH-S $\rightleftharpoons$ NH-S	
	1	2	1	2
$E_{\text{TS}}$	-738.569175	-734.310145	-738.568810	-734.308541
$E_{\text{forward}}$	51.64	99.18	42.18	92.55
$E_{\text{reverse}}$	58.60	105.47	61.04	111.72
$\Delta G_{298}^{\circ}$	-7.21	-7.85	-19.81	-26.32
$\Delta H_{298}^{\circ}$	-6.12	-5.44	-18.13	-15.55
$T\Delta S_{298}^{\circ}$	1.09	2.41	1.68	10.77
$K_{298}^{\#}$ , forward	$6.80 \times 10^{-8}$	$8.88 \times 10^{-18}$	$2.06 \times 10^{-6}$	$4.83 \times 10^{-15}$
$K_{298}^{\#}$ , reverse	$6.80 \times 10^{-8}$	$3.00 \times 10^{-9}$	$6.95 \times 10^{-10}$	$1.18 \times 10^{-19}$
$k_{298}$ , forward	$4.22 \times 10^5$	$5.51 \times 10^{-5}$	$1.28 \times 10^7$	$3.00 \times 10^{-2}$
$k_{298}$ , reverse	$0.19 \times 10^5$	$0.23 \times 10^{-5}$	$4.31 \times 10^3$	$7.32 \times 10^{-7}$
	OH-R1 $\rightleftharpoons$ NH-R1		OH-S1 $\rightleftharpoons$ NH-S1	
	2	1	2	
$E_{\text{TS}}$	-734.336765	-738.590286	-734.335564	
$E_{\text{forward}}$	56.62	10.38	41.80	
$E_{\text{reverse}}$	69.03	27.70	60.18	
$\Delta G_{298}^{\circ}$	-10.56	-15.51	-16.86	
$\Delta H_{298}^{\circ}$	-12.01	-16.09	-17.96	
$T\Delta S_{298}^{\circ}$	-1.45	-0.58	-1.10	
$K_{298}^{\#}$ , forward	$1.30 \times 10^{-11}$	0.78	$1.16 \times 10^{-6}$	
$K_{298}^{\#}$ , reverse	$1.84 \times 10^{-13}$	$1.49 \times 10^{-3}$	$1.00 \times 10^{-9}$	
$k_{298}$ , forward	80.73	$4.84 \times 10^{12}$	$7.20 \times 10^6$	
$k_{298}$ , reverse	1.14	$9.25 \times 10^9$	$7.97 \times 10^3$	

1: B3LYP/6-31+G(d); 2: HF/6-31+G(d)



intramolecular proton transfers in guanine and *isoC* in the gas phase [6, 14].

Obviously, the rate constants and the activation equilibrium constants are very sensitive to the theoretical method used. Nevertheless, the forward reactions have higher rate constants than the reverse ones. Therefore, the NH-supersystems should be much more favorable than the OH- ones.

## Conclusion

In this paper results are presented for the theoretical study (B3LYP and HF with the 6-31+G(d) basis set) of the structural, thermodynamic and vibrational parameters of eight supersystems composed between *R/S*-lactic acid and two most stable tautomeric forms of isocytosine (an important residue in the guanine skeleton). Proton-transfer phenomena assisted by molecules of lactic acid were also studied. Several principal conclusions should be mentioned:

1. According to the thermodynamic data, the NH-R1 supersystem should be the most stable. The OH-S supersystem is the least stable aggregate with the largest, and the lowest equilibrium constant of the complexation reaction.
2. All supersystems formed between the OH group of *R/S*-LA have higher entropy factors compared to the absolute value of the enthalpy. In other words, the steric effects dominate during all complexations. On the other hand, the complexations between the carboxylic group and *isoC* are accompanied by larger energy (enthalpy) changes than steric (entropy) alterations. These complexations have a negative variation of the free Gibbs energy, which means that they are thermodynamically favored, and the equilibrium is shifted towards the supersystem.
3. The bonding analysis showed that the bonding energies and the interaction energies of the supersystems are larger for the complexes constituted between the carboxylic function of the *R/S*-LA and *isoC*. According to these parameters, all complexes of the NH-type are the most stable. The complexations with the OH-form of *isoC* lead to the supersystems with lower stability.
4. The estimated energy barriers of the tautomeric conversions showed that the intermolecular H-exchange between carboxylic group and *isoC* has several times lower energy barriers than the H-exchange including the OH group of *R/S*-LA. As a whole, the energy barriers estimated here are several times lower than those found for the intramolecular proton transfers in guanine and *isoC* in the gas phase [6, 13, 14]. Comparing the rate constants and the energy barriers, one can see that the transformation OH-S1NH-S1 is the most favorable. In other words, *S*-LA facilitates the proton exchange processes in *isoC* to a greater extent.

**Acknowledgements** One of the authors (VBD) thanks the fund "Scientific Researches" at the University of Plovdiv. The authors thank the Central Services of ZID TU Vienna, where the Gaussian 98 calculations were performed.

## References

1. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2002) *Phys Chem Chem Phys* 4:740–750
2. Mons M, Dimicoli I, Piuze F, Tardivel B, Elhanine M (2002) *J Phys Chem A* 106:5088–5094
3. Chin W, Mons M, Dimicoli I, Piuze F, Tardivel B, Elhanine M (2002) *Eur Phys J D* 20:347–355
4. Piuze F, Mons M, Dimicoli I, Tardivel B, Zhao Q (2001) *Chem Phys* 270:205–214
5. Shishkin OV, Sukhanov OS, Gorb L, Leszczynski J (2002) *Phys Chem Chem Phys* 4:5359–5364
6. Delchev VB, Mikosch H (2004) *Scientific works. Scientists Union of Bulgaria—Plovdiv, IV*:130–135
7. Gorb L, Leszczynski J (1998) *J Am Chem Soc* 120:5024–5032
8. Felice RD, Calzolari A, Molinari E (2001) *Phys Review B* 65:045104/1–045104/10
9. Meyer M, Brandl M, Sühnel J (2001) *J Chem Phys A* 105:8223–8225
10. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2002) *Phys Chem Chem Phys* 4:732–739
11. Kwiatkowski JS, Leszczynski J (1990) *J Mol Str (THEOCHEM)* 208:35–44
12. Burchenal JH, Ciovacco K, Kalaher K (1976) *Cancer Res* 36:1520–1523
13. Gorb J, Podolyan Y, Leszczynski J (1999) *J Mol Str (THEOCHEM)* 487:47–55
14. Gorb L, Leszczynski J (1998) *J Am Chem Soc* 112:5024–5032
15. Gorb L, Leszczynski J (1997) *Int J Quantum Chem* 65:759–765
16. Gorb L, Leszczynski J (1998) *Int J Quantum Chem* 70:855–862
17. Shishkin OV (1998) *J Mol Str* 447:1–5
18. Riggs NV (1991) *Chem Phys Lett* 177:447–450
19. Leszczynski J (1998) *J Phys Chem A* 102:2357–2362
20. Sponer J, Hobza P (1994) *J Mol Str (THEOCHEM)* 304:35–40
21. Sponer J, Hobza P (1994) *J Phys Chem* 98:3161–3164
22. Podolyan Y, Gorb L, Leszczynski J (2003) *Int J Mol Sci* 4:410–421
23. Langer H, Doltsinis NL (2003) *J Chem Phys* 118:5400–5407
24. Mons M, Dimicoli I, Piuze F (2002) *Int Rev Phys Chem* 21:101–135
25. Sevilla MD, Becker D, Yan M, Summerfield SR (1991) *J Phys Chem* 95:3409–3415
26. Hendricks JH, Lyapustina SA, de Clercq HL, Bowen KH (1998) *J Chem Phys* 108:8–11
27. Orlov VM, Smirnov AN, Varshavsky YM (1976) *Tetrahedron Lett* 48:4377–4378
28. Li X, Cai Z, Sevilla MD (2001) *J Phys Chem B* 105:10115–10123
29. 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 O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson JA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez G, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (1998) *Gaussian 98, Revision A.3*. Gaussian Inc, Pittsburgh, Pennsylvania
30. Atkins PW (1994) In: *Physical Chemistry*. Oxford University Press, pp 147–170

31. Moor WJ (1974) In: *Physical Chemistry*, Longman, pp 282–299
32. Boys SF, Bernadi F (1970) *Mol Phys* 19:553–566
33. Hobza P, Zahradnik R (1989) In: *Intermolecular Complexes*. Mir, Moscow, pp 36
34. Meunier A, Levy B, Bertier G (1973) *Theor Chim Acta* 29:49–52
35. Mayer I, Surjan PR (1992) *Chem Phys Lett* 191:497–499
36. Turi L, Dannenberg JJ (1993) *J Phys Chem* 97:2488–2490
37. Simon S, Duran M, Dannenberg JJ (1999) *J Phys Chem A* 103:1640–1643
38. Cleland WW, Kreevoy MM (1994) *Science* 103:264–269
39. Delchev VB, Mikosch H (2004) *Monatsh Chem* 135:1373–1387
40. Chandra AK, Nguyen MT, Uchimaru T, Zeegers-Huyskens (1999) *J Phys Chem* 103:8853–8860
41. van Mourik T, Price SL, Clary DC (1999) *J Phys Chem A* 103:1611–1618
42. Li X, Cai Z, Sevilla MD (2002) *J Phys Chem A* 106:9345–9351
43. Barsky D, Colvin ME (2000) *J Phys Chem A* 104:8570–8576
44. Foresman JB, Frisch AE (1995/96) In: *Exploring Chemistry with Electronic Structure Methods*, 2nd edn. Gaussian Inc, Pittsburg, pp 99
45. SDBSWeb: <http://www.aist.go.jp/RIODB/SDBS/>