

**STABILITY, POLARITY, INTRAMOLECULAR INTERACTIONS AND  $\pi$ -ELECTRON DELOCALIZATION FOR ALL EIGHTEEN TAUTOMERS ROTAMERS OF URACIL. DFT STUDIES IN THE GAS PHASE**

Ewa D. RACZYŃSKA<sup>a,\*</sup>, Katarzyna ZIENTARA<sup>b</sup>, Tomasz M. STĘPNIEWSKI<sup>b</sup> and Katarzyna KOLCZYŃSKA<sup>b</sup>

<sup>a</sup> Department of Chemistry, Warsaw University of Life Sciences, 02-776 Warszawa, Poland; e-mail: ewa\_raczynska@sggw.pl

<sup>b</sup> Interdisciplinary Department of Biotechnology, Warsaw University of Life Sciences, 02-726 Warsaw, Poland

Received June 27, 2008

Accepted September 3, 2008

Published online January 22, 2009

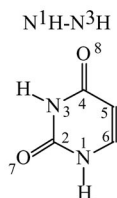
*Dedicated to the memory of Professor Otto Exner.*

Complete tautomeric equilibria were investigated for uracil at the DFT(B3LYP)/6-311+G(d,p) level to establish the stability order of all possible 18 tautomers-rotamers in the gas phase and to characterize their internal effects, polarity and aromatic character. Although the di-NH form strongly predominates (100%) in the mixture, the NH-OH, di-OH and CH-NH forms can be also considered. The favored tautomer is moderately stabilized by intramolecular interactions (attractions of the NH and C=O groups); it is also moderately polar and moderately delocalized. Stability of the functional groups (both amide functions) seems to be more important than intramolecular interactions, polarity and aromaticity. This is probably the main factor that dictates the tautomeric preferences in the uracil molecule.

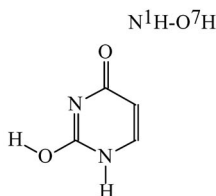
**Keywords:** Uracil; Tautomers-rotamers;  $\pi$ -Electron delocalization; DFT.

It is well known that prototropic tautomerism occurring in nucleobases, pyrimidines (uracil, thymine and cytosine) and purines (adenine and guanine), is one of the phenomena responsible for spontaneous point mutations of DNA and RNA<sup>1</sup>. During tautomeric interconversion, it is sufficient that a proton is transferred from one to another conjugated site in nucleobases and their capability of H-bonding dramatically changes. Although tautomeric equilibria for nucleobases have been extensively investigated by both experimental and theoretical methods, the CH tautomers have been frequently omitted. In the case of uracil, usually six tautomers (di-NH, NH-OH, and di-OH forms) have been considered in which two pro-

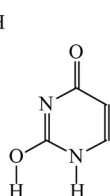
tons move between four heteroatoms (two *exo* oxygens and two *endo* nitrogens). When rotational isomerism of the *exo*-OH group has been taken into account, the structures and stabilities of thirteen tautomers-rotamers have been characterized<sup>2</sup>. However, the proton can be also intramolecularly transferred from O and/or N to C5 atom. All tautomeric interconversions lead to nine tautomers for uracil. When rotational isomerism of the *exo*-OH group is considered, eighteen tautomers-rotamers are possible (Scheme 1). Some CH tautomers have been recently considered for valence-type and dipole-bonded anionic states of uracil<sup>3</sup>. However,  $\pi$ -electron delocalization has been discussed solely for a few selected NH and OH tautomers<sup>4,5</sup>.



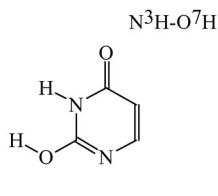
U1



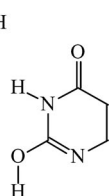
U2a



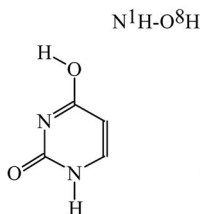
U2b



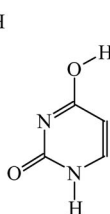
U3a



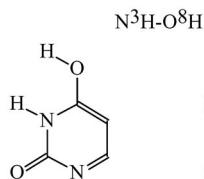
U3b



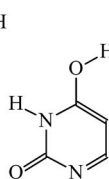
U4a



U4b



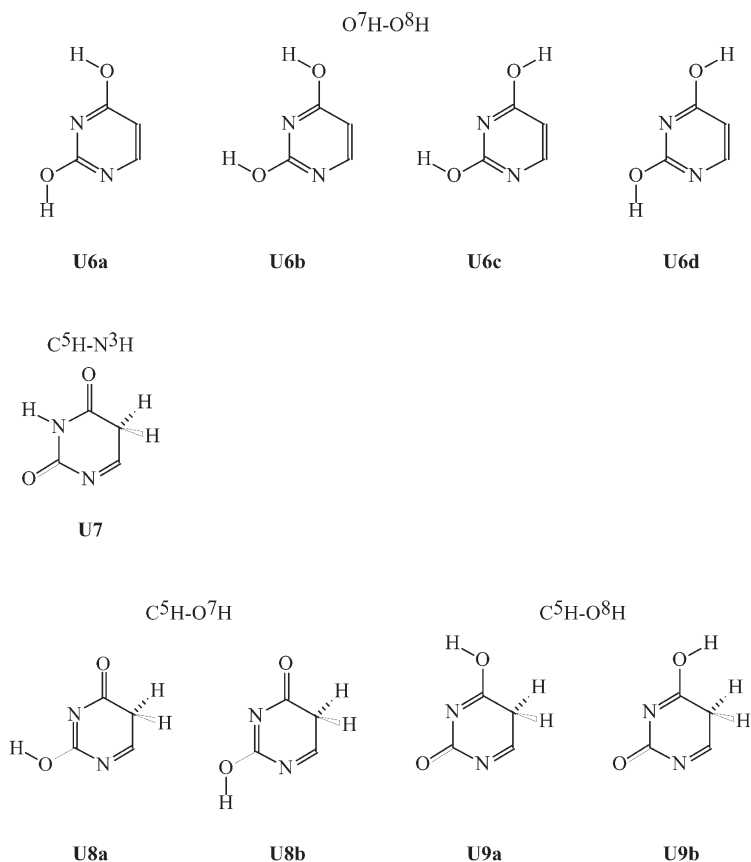
U5a



U5b

Scheme 1 continued

Understanding the chemistry and biochemistry of uracil, particularly various processes such as substitution and oxidation reactions, point mutations developing in RNA and DNA replications, etc., require understanding its structure, intramolecular conversions and intramolecular interactions. Since proton-transfer reactions are very fast and separation of individual tautomers is very difficult, experiments performed for tautomeric systems, particularly those in the gas phase and solution, do not give complete structural picture. Very often one can measure macroscopic parameters for tautomeric mixtures but not for individual tautomers. Moreover, when a single form dominates, it is difficult to detect the rare tautomers. Most



SCHEME 1  
Possible tautomers-rotamers for uracil

spectroscopic methods are incapable to detect less than 0.1% of a minor tautomer. Fortunately, structural investigations can be carried out using various computational techniques. These techniques permit to study separately all possible tautomeric forms, all possible tautomeric conversions, and all possible internal effects. Such information is essential for estimation of microscopic parameters which are fundamental for particular processes.

The aim of the present work is to study all possible tautomeric equilibria for neutral uracil in the gas phase, i.e., the intramolecular proton-transfer from N to O (NH/OH), from O to O (OH/OH), from N to N (NH/NH), from O to C (OH/CH), and from N to C atom (NH/CH). These studies give the possibilities to characterize all the eighteen possible tautomers-rotamers of uracil, to establish their stability order, to analyse internal effects, to estimate their polarity and  $\pi$ -electron delocalization. Of numerous DFT methods<sup>6</sup>, we chose the Becke three-parameter hybrid exchange functional with the non-local correlation functional of Lee, Yang and Parr<sup>7</sup> (B3LYP), which is sufficient to study the proton-transfer reactions for tautomerizing and *ortho*-substituted systems<sup>2-4,8,9</sup>. Since various intramolecular interactions may take place in uracil tautomers-rotamers (e.g. repulsions of the lone pairs of electrons of the neighboring heteroatoms (O and N), repulsions of the neighboring OH, NH and CH groups, intramolecular interactions between the neighboring functional groups such as OH...N(-N=), NH...O(C=O), and NH...O(OH) interactions), the 6-311+G(d,p) basis set<sup>10</sup> with the diffuse and polarization functions was applied. To determine distribution of  $\pi$ -electrons for all tautomers-rotamers of uracil, the geometry-based HOMA (harmonic oscillator model of aromaticity) procedure<sup>11-13</sup> and the CX (X = C, N, O) bond lengths of simple reference compounds were applied to geometries of all 18 possible isomers of uracil optimized at the B3LYP/6-311+G(d,p) level.

## COMPUTATIONAL DETAILS

Geometries of all the eighteen possible tautomers-rotamers of neutral uracil (Scheme 1) were fully optimized without symmetry constraints using the DFT(B3LYP) method<sup>7</sup> and the 6-311+G(d,p) basis set<sup>10</sup>. For all minima (with real frequencies), thermodynamic parameters such as electronic energy ( $E$ ), enthalpy ( $H = E + pV$ ), entropy ( $S$ ), and Gibbs (free) energies ( $G = H - TS$  for  $T = 298.15$  K) were calculated using the same level of theory. For tautomeric conversions, the relative Gibbs energy ( $\Delta G = G(\mathbf{U}_i) - G(\mathbf{U}_j)$ ) and the tautomeric equilibrium constants ( $pK_T = \Delta G/2.303RT$ ) were calculated and, then, the percentage contents of all tautomers-rotamers were estimated for the

tautomeric mixture of gaseous uracil. The  $\Delta G$  values include the changes in electronic energy, in zero-point energy (ZPE) and in thermal corrections to energy and entropy (vibrational, rotational, and translational).

For calculation of the HOMA index, the same equation was used as that proposed for the reformulated HOMA index<sup>13</sup>:

$$\text{HOMA} = \left[ 1 - \alpha_{\text{CX}} \sum_{i=1}^n \{R_{\text{opt}}(\text{CX}) - R_i(\text{CX})\}^2 \right] / n$$

where  $n$  is the number of bonds taken into account,  $\alpha_{\text{CX}}$  is a normalization constant:  $\alpha_{\text{CX}} = 2\{[R_s(\text{CX}) - R_{\text{opt}}(\text{CX})]^2 + [R_d(\text{CX}) - R_{\text{opt}}(\text{CX})]^2\}^{-1}$ ,  $R_s(\text{CX})$  and  $R_d(\text{CX})$  are the reference CX ( $X = \text{C}, \text{N}, \text{O}$ ) single and double bonds,  $R_{\text{opt}}(\text{CX})$  is the optimum CX bond length (assumed to be realized when full delocalization of  $\pi$ -electrons takes place), and  $R_i(\text{CX})$  are the running bond lengths in the system. The following  $R_s(\text{CX})$ ,  $R_d(\text{CX})$  and  $R_{\text{opt}}(\text{CX})$  values (in Å), calculated at the B3LYP/6-311+G(d,p) level for simple model compounds, were taken here: 1.536 (cyclohexane), 1.335 (cyclohexene) and 1.394 (benzene) for CC bonds, 1.465 (piperidine), 1.269 (2,3,4,5-tetrahydropyridine) and 1.334 (1,3,5-triazine) for CN bonds, and 1.433 (cyclohexanol), 1.213 (cyclohexanone) and 1.249 (carbonate) for CO bonds. These  $R$  values lead to the following normalization constant  $\alpha_{\text{CX}}$ : 84.58, 93.52 and 56.90 for the CC, CN and CO bond, respectively. All calculations were performed using the Gaussian 03 program<sup>14</sup>.

## RESULTS AND DISCUSSION

### Stability

For calculations performed at the DFT(B3LYP)/6-311+G(d,p) level, eighteen isomers of uracil (Scheme 1) were considered, i.e., one  $\text{N}^1\text{H}-\text{N}^3\text{H}$  tautomer (**U1**), two rotamers of the  $\text{N}^1\text{H}-\text{O}^7\text{H}$  tautomer (**U2a** and **U2b**), two rotamers of the  $\text{N}^3\text{H}-\text{O}^7\text{H}$  tautomer (**U3a** and **U3b**), two rotamers of the  $\text{N}^1\text{H}-\text{O}^8\text{H}$  tautomer (**U4a** and **U4b**), two rotamers of the  $\text{N}^3\text{H}-\text{O}^8\text{H}$  tautomer (**U5a** and **U5b**), four rotamers of the  $\text{O}^7\text{H}-\text{O}^8\text{H}$  tautomer (**U6a-U6d**), one  $\text{C}^5\text{H}-\text{N}^3\text{H}$  tautomer (**U7**), two rotamers of the  $\text{C}^5\text{H}-\text{O}^7\text{H}$  tautomer (**U8a** and **U8b**), and two rotamers of the  $\text{C}^5\text{H}-\text{O}^8\text{H}$  tautomer (**U9a** and **U9b**). For all the isomers, the minima with real frequencies were found. The **U1** tautomer (called also the lactam form) containing two amide groups in the ring has the lowest energy. As shown in numerous experimental and theoretical papers, this

form is the most stable in the gas phase as well as in solution and in the solid state<sup>2-5,15</sup>.

The DFT enthalpies ( $\Delta H$ ), Gibbs energies ( $\Delta G$ ) and entropy terms ( $T\Delta S$ ) of uracil isomers relative to those of **U1** are summarized in Table I. As it could be expected, an application of larger basis sets has no important effect on the relative stabilities. No isomer has the  $G$  value close to that of **U1**. The  $\Delta G$  values of ten isomers are larger than 10 kcal mol<sup>-1</sup> and lower than 20 kcal mol<sup>-1</sup>. Three isomers possess the  $\Delta G$  values between 20 and 30 kcal mol<sup>-1</sup>, and three isomers have the  $\Delta G$  values between 30 and 40 kcal mol<sup>-1</sup>. For one isomer, the  $\Delta G$  value is larger than 40 kcal mol<sup>-1</sup>. The order of decreasing DFT stability for the eighteen uracil isomers is as fol-

TABLE I

Relative thermodynamic parameters ( $\Delta H$ ,  $T\Delta S$  and  $\Delta G$ )<sup>a</sup>, percentage contents (% **Ui**) and dipole moments ( $\mu$  in debye) estimated at the B3LYP/6-311+G(d,p) level for all 18 possible tautomers-rotamers of uracil

Isomer	$\Delta H^a$	$T\Delta S^a$	$\Delta G^a$	% <b>Ui</b>	$\mu^b$ , D
<b>U1</b>	0.0	0.0	0.0	100	4.58
<b>U2a</b>	19.5	0.1	19.4	$<6 \times 10^{-13}$	6.60
<b>U2b</b>	29.9	0.7	29.1	$<4 \times 10^{-20}$	9.16
<b>U3a</b>	19.3	0.3	19.0	$<1 \times 10^{-12}$	2.31
<b>U3b</b>	11.4	-0.2	11.6	$<3 \times 10^{-7}$	3.33
<b>U4a</b>	12.1	-0.1	12.3	$<1 \times 10^{-7}$	4.99
<b>U4b</b>	19.1	0.1	19.0	$<1 \times 10^{-12}$	7.83
<b>U5a</b>	24.1	0.4	23.7	$<4 \times 10^{-16}$	5.98
<b>U5b</b>	21.3	0.1	21.2	$<3 \times 10^{-14}$	7.36
<b>U6a</b>	13.6	-0.4	13.9	$<6 \times 10^{-9}$	1.32
<b>U6b</b>	14.7	-0.3	15.0	$<1 \times 10^{-9}$	2.47
<b>U6c</b>	19.0	-0.3	19.2	$<8 \times 10^{-13}$	3.81
<b>U6d</b>	18.9	-0.3	19.2	$<9 \times 10^{-13}$	4.19
<b>U7</b>	17.5	1.0	16.6	$<7 \times 10^{-11}$	3.75
<b>U8a</b>	34.6	0.5	34.2	$<9 \times 10^{-24}$	3.16
<b>U8b</b>	42.1	0.5	41.7	$<3 \times 10^{-29}$	5.98
<b>U9a</b>	34.2	0.5	33.7	$<2 \times 10^{-23}$	6.01
<b>U9b</b>	35.8	0.4	35.3	$<1 \times 10^{-24}$	8.29

<sup>a</sup>  $\Delta P = P(\mathbf{Ui}) - P(\mathbf{U1})$ , where  $P$  is  $H$ ,  $TS$  or  $G$ . All in kcal mol<sup>-1</sup> at 298.15 K.

lows: **U1**, **U3b**, **U4a**, **U6a**, **U6b**, **U7**, **U4b**, **U3a**, **U6d**, **U6c**, **U2a**, **U5b**, **U5a**, **U2b**, **U8a**, **U9a**, **U8b**, and **U9b**. Jalbout et al.<sup>2a</sup> reported almost the same order of decreasing stability for the thirteen NH and OH tautomers at the DFT(B3LYP)/6-31+G(d,p) level: **U1**, **U3b**, **U4a**, **U6a**, **U6b**, **U6d**, **U6c**, **U4b**, **U3a**, **U2a**, **U5b**, **U5a**, and **U2b**. They have not considered the five CH tautomers. Using the MP2 and CCSD(T) methods and the same basis set, they found slightly different stability order for rare tautomers. However, independently of the level of calculations, the **U1** form has lower Gibbs energy than the other ones by more than 10 kcal mol<sup>-1</sup>.

All computations indicate that the tautomeric mixture of uracil consists mainly of the **U1** tautomer (100%). The amounts of other forms are lower than 1 ppm. Interestingly, the percentage content of the **U7** tautomer (C<sup>5</sup>H–N<sup>3</sup>H form) containing the C5-sp<sup>3</sup> atom in the ring is close to those of the di-OH isomers (**U6a**, **U6b**, **U6c**, and **U6d**). This means that the **U7** tautomer cannot be neglected in the tautomeric mixture of uracil because it may be responsible for direction of some substitution, oxidation or degradation reactions, similarly as di-OH isomers. Other CH tautomers (**U8a**, **U9a**, **U8b**, and **U9b**) are less important. Their percentage contents are lower than 10<sup>-20</sup>%.

The relative entropy terms ( $T\Delta S$ ) are not larger than  $\pm 1$  kcal mol<sup>-1</sup>. This indicates that tautomeric interconversions for uracil are almost isoentropic in the gas phase<sup>16</sup>. Generally, there are not large structural changes in tautomerization. Some exceptions are those being a consequence of the rotational isomerism of the *ex*-OH group for the OH forms and those resulting from loss of the ring planarity for the CH forms.

### Polarity

Polarity of each uracil isomer may be described by its dipole moment ( $\mu$ ) predicted at the same level of theory as its thermodynamic parameters. A first perusal of the calculated  $\mu$  values indicates that the favored tautomer **U1** is moderately polar (Table I). Its  $\mu$  value is close to that found experimentally<sup>17</sup>.

The rare CH–NH tautomer **U7** has only slightly lower  $\mu$  value (by 0.08 D) than the favored di-NH tautomer **U1**. One (**U6a**) of the rare di-OH isomers has the smallest polarity ( $\mu \approx 1$  D). One (**U2b**) of the rare NH–OH isomers has the greatest polarity ( $\mu \approx 9$  D). These DFT predictions show that the increasing polarity order (**U6a**, **U3a**, **U6b**, **U8a**, **U3b**, **U7**, **U6c**, **U6d**, **U1**, **U4a**, **U5a**, **U8b**, **U9a**, **U2a**, **U5b**, **U9b**, and **U2b**) is not parallel to the decreasing stability order (**U1**, **U3b**, **U4a**, **U6a**, **U6b**, **U7**, **U4b**, **U3a**, **U6d**, **U6c**, **U2a**,

**U5b, U5a, U2b, U8a, U9a, U8b, and U9b**). Hence, polarity cannot be considered as the main factor that influences the tautomeric preference for uracil as it has been suggested for hydroxypyridines<sup>18</sup>.

### Rotation of the *exo*-OH Group

Due to rotational isomerism of the *exo*-OH group in the NH-OH (**U2–U4**) and CH-OH (**U8–U9**) uracil tautomers, two conformations are possible (Scheme 1), one with the hydroxy H-atom synperiplanar (**a**) and the other one with the hydroxy H-atom antiperiplanar (**b**) to the N<sup>3</sup> atom. This conformational difference leads to difference in stabilities of the **a** and **b** isomers (Table II). Similarly as for hydroxypyridines<sup>18,19</sup>, an intramolecular H-bond between the *exo*-OH and *endo*-N(=N) group strongly stabilizes **U2a** (O<sup>7</sup>H...N<sup>3</sup>), **U3b** (O<sup>7</sup>H...N<sup>1</sup>), **U4a** (O<sup>8</sup>H...N<sup>3</sup>), and **U9a** (O<sup>8</sup>H...N<sup>3</sup>). The **U2a**, **U3b** and **U4a** isomers are additionally stabilized by interactions of the *endo*-NH and *exo*-O(OH or C=O) groups, i.e., by the N<sup>1</sup>H...O<sup>7</sup>H, N<sup>3</sup>H...O<sup>7</sup>H and N<sup>1</sup>H...O<sup>7</sup>=C<sup>2</sup> interactions, respectively. On the other hand, repulsion(s) between the lone electron pairs of the neighboring heteroatoms and/or the neighboring XH groups (X = O, N, C), destabilize **U2b** (N<sup>1</sup>H–O<sup>7</sup>H and O<sup>7</sup>–N<sup>3</sup>), **U3a** (N<sup>3</sup>H–O<sup>7</sup>H and O<sup>7</sup>–N<sup>1</sup>), **U4b** (C<sup>5</sup>H–O<sup>8</sup>H and N<sup>3</sup>–O<sup>8</sup>), and **U9b** (C<sup>5</sup>H–O<sup>8</sup>H, N<sup>1</sup>–O<sup>8</sup> and N<sup>3</sup>–O<sup>8</sup>). Hence, differences between the Gibbs energies ( $\Delta G$ ) of the **a** and **b** isomers are exceptionally high (7–10 kcal mol<sup>-1</sup>).

The  $\Delta G$  value for the conversion **U5a**→**U5b** is considerably smaller (2.5 kcal mol<sup>-1</sup>), because the isomers differ only by the repulsion of the neighboring XH groups (N<sup>3</sup>H–O<sup>8</sup>H and C<sup>5</sup>H–O<sup>8</sup>H) and by the XH...O(OH) interactions (C<sup>5</sup>H...O<sup>8</sup>H and N<sup>3</sup>H...O<sup>8</sup>H). The  $\Delta G$  value is also small for the

TABLE II  
Relative Gibbs (free) energies ( $\Delta G$ )<sup>a</sup> estimated at the DFT(B3LYP)/6-311+G(d,p) level for the OH rotamers of uracil

Pair of rotamers	$\Delta G^a$	Pair of rotamers	$\Delta G^a$
<b>U2a</b> → <b>U2b</b>	-9.7	<b>U6a</b> → <b>U6d</b>	-5.2
<b>U3a</b> → <b>U3b</b>	7.5	<b>U6b</b> → <b>U6c</b>	-4.2
<b>U4a</b> → <b>U4b</b>	-6.7	<b>U6c</b> → <b>U6d</b>	0.0
<b>U5a</b> → <b>U5b</b>	2.5	<b>U8a</b> → <b>U8b</b>	-1.6
<b>U6a</b> → <b>U6b</b>	-1.1	<b>U9a</b> → <b>U9b</b>	-7.6

<sup>a</sup> In kcal mol<sup>-1</sup> at 298.15 K.

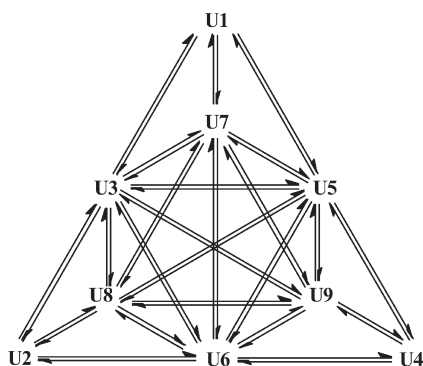


conversion **U8a**→**U8b** ( $-1.6 \text{ kcal mol}^{-1}$ ) because the isomers differ only by the repulsion of the lone-electron pairs of the neighboring heteroatoms ( $\text{N}^1\text{-O}^7$  and  $\text{N}^3\text{-O}^7$ ). Both isomers are stabilized by the  $\text{OH}\cdots\text{N}(-\text{N}=\)$  interactions ( $\text{O}^7\text{H}\cdots\text{N}^3$  and  $\text{O}^7\text{H}\cdots\text{N}^1$ ).

For the di-OH tautomer **U6** containing two *exo*-OH groups, rotational isomerism of these groups leads to four conformations (**a**–**d**). All of them are stabilized by the intramolecular  $\text{OH}\cdots\text{N}(-\text{N}=\)$  interactions. Since two H-bonds are possible for **U6a** ( $\text{O}^7\text{H}\cdots\text{N}^1$  and  $\text{O}^8\text{H}\cdots\text{N}^3$ ) and **U6b** ( $\text{O}^7\text{H}\cdots\text{N}^3$  and  $\text{O}^8\text{H}\cdots\text{N}^3$ ), they have lower energies than **U6c** and **U6d**, for which only one H-bond can be formed ( $\text{O}^7\text{H}\cdots\text{N}^3$  and  $\text{O}^7\text{H}\cdots\text{N}^1$ , respectively). Therefore, the  $\Delta G$  values for the conversion **U6a**→**U6d** and **U6b**→**U6c** are larger ( $4\text{--}5 \text{ kcal mol}^{-1}$ ) than those for **U6a**→**U6b** and **U6c**→**U6d** ( $0\text{--}1 \text{ kcal mol}^{-1}$ ). All these observations show how important are intramolecular interactions between functional groups for rare tautomers-rotamers of uracil<sup>2a,2b</sup>. Any interactions with external species, which destroy the intramolecular interactions, may change the energetic stability order of uracil isomers and, consequently, change the proton-transfer reactions.

### Tautomeric Equilibria

For all eighteen tautomers-rotamers of uracil, 153 equilibria are possible, 141 intramolecular proton-transfers in tautomeric conversions and 12 equilibria being a consequence of rotational isomerism of the *exo*-OH group(s). Selected tautomeric equilibria are given in Scheme 2. The proton may be intramolecularly transferred from N to O, from N to N, from O to O, from



SCHEME 2  
Selected tautomeric equilibria for uracil

N to C, and from O to C by the 1,3-, 1,5- or 1,7-proton shift. The DFT Gibbs energies ( $\Delta G$ ) and tautomeric equilibrium constants ( $pK_T$ ) for selected tautomeric conversions are listed in Table III.

TABLE III  
Gibbs energies ( $\Delta G$ )<sup>a</sup> and tautomeric equilibrium constants ( $pK_T$ )<sup>b</sup> estimated at the DFT(B3LYP)/6-311+G(d,p) level for selected tautomeric conversions

Proton transfer	Proton shift	Tautomer pair	$\Delta G^a$	$pK_T^b$
N <sup>1</sup> H→O <sup>7</sup> H	1,3	<b>U1→U3b</b>	11.6	8.5
N <sup>3</sup> H→O <sup>7</sup> H	1,3	<b>U1→U2a</b>	19.4	14.2
N <sup>3</sup> H→O <sup>8</sup> H	1,3	<b>U1→U4a</b>	12.3	9.0
N <sup>1</sup> H→O <sup>8</sup> H	1,5	<b>U1→U5b</b>	21.2	15.5
N <sup>1</sup> H→O <sup>7</sup> H	1,3	<b>U4a→U6a</b>	1.7	1.2
N <sup>3</sup> H→O <sup>7</sup> H	1,3	<b>U5a→U6b</b>	-8.7	-6.4
N <sup>3</sup> H→O <sup>8</sup> H	1,3	<b>U3b→U6a</b>	2.4	1.7
N <sup>1</sup> H→O <sup>8</sup> H	1,5	<b>U2a→U6b</b>	-4.4	-3.2
N <sup>1</sup> H→N <sup>3</sup> H	1,3	<b>U2a→U3a</b>	-0.4	-0.3
N <sup>1</sup> H→N <sup>3</sup> H	1,3	<b>U4b→U5b</b>	2.2	1.6
N <sup>1</sup> H→N <sup>3</sup> H	1,3	<b>U2b→U3b</b>	-7.6	-12.9
N <sup>1</sup> H→N <sup>3</sup> H	1,3	<b>U4a→U5a</b>	11.4	8.4
O <sup>7</sup> H→O <sup>8</sup> H	1,5	<b>U2a→U4a</b>	-7.1	-5.2
O <sup>7</sup> H→O <sup>8</sup> H	1,7	<b>U3b→U5b</b>	9.6	7.1
O <sup>7</sup> H→O <sup>8</sup> H	1,7	<b>U2b→U4b</b>	-10.1	-7.4
O <sup>7</sup> H→O <sup>8</sup> H	1,7	<b>U3a→U5a</b>	4.7	3.4
N <sup>1</sup> H→C <sup>5</sup> H	1,3	<b>U1→U7</b>	16.6	12.1
N <sup>1</sup> H→C <sup>5</sup> H	1,3	<b>U2a→U8a</b>	14.3	10.5
N <sup>1</sup> H→C <sup>5</sup> H	1,3	<b>U4a→U9a</b>	21.9	16.1
N <sup>3</sup> H→C <sup>5</sup> H	1,3	<b>U5a→U9a</b>	10.5	7.7
N <sup>3</sup> H→C <sup>5</sup> H	1,5	<b>U3b→U8b</b>	23.8	17.4
O <sup>7</sup> H→C <sup>5</sup> H	1,5	<b>U3b→U7</b>	5.0	3.7
O <sup>7</sup> H→C <sup>5</sup> H	1,5	<b>U6a→U9a</b>	20.2	14.8
O <sup>8</sup> H→C <sup>5</sup> H	1,3	<b>U5b→U7</b>	-16.6	-12.1
O <sup>8</sup> H→C <sup>5</sup> H	1,7	<b>U6a→U8b</b>	21.4	15.7

<sup>a</sup> In kcal mol<sup>-1</sup> at 298.15 K. <sup>b</sup> At 298.15 K.

The  $pK_T$  values for the favored proton-transfers from the N to O atom,  $N^1H \rightarrow O^7H$  for **U1**  $\rightarrow$  **U3b** and  $N^3H \rightarrow O^8H$  for **U1**  $\rightarrow$  **U4a**, are almost twice smaller than those for the other ones,  $N^1H \rightarrow O^8H$  for **U1**  $\rightarrow$  **U5b** and  $N^3H \rightarrow O^7H$  for **U1**  $\rightarrow$  **U2a**. This suggests that one of the amide groups of **U1** ( $HN^1-C^2=O^7$  or  $HN^3-C^4=O^8$ ) can be easier converted to the hydroxy-imino group ( $N^1=C^2-O^7H$  or  $N^3=C^4-O^8H$ ) giving **U3b** or **U4a**, respectively, than the two amide groups of **U1** ( $HN^1-C^2=O^7$  and  $HN^3-C^4=O^8$ ) to the amide ( $O^7=C^2-N^3H$ ) and hydroxy-enimino ( $N^1=C^6-C^5=C^4-O^8H$ ) groups in **U5b**, or to the hydroxy-imino ( $HO^7-C^2=N^3$ ) and enamino-oxo ( $HN^1-C^6=C^5-C^4=O^8$ ) groups in **U2a**. The  $pK_T$  values for the successive proton-transfers from the N to O atom in the NH-OH isomers having lower energies,  $N^1H \rightarrow O^7H$  for **U4a**  $\rightarrow$  **U6a** and  $N^3H \rightarrow O^8H$  for **U3b**  $\rightarrow$  **U6a**, are very small ( $pK_T$  close to 1–2) indicating that any changes of atom environment may influence these equilibria. The successive proton-transfers from the N to O atom for the other NH-OH isomers,  $N^1H \rightarrow O^8H$  for **U2a**  $\rightarrow$  **U6b** and  $N^3H \rightarrow O^7H$  for **U5a**  $\rightarrow$  **U6b**, are very favorable because the di-OH isomer **U6b** has smaller energy than the NH-OH isomers **U2a** and **U5a**.

For the NH-OH isomers, the proton may be also transferred between atoms of the same element, i.e., from N to N or from O to O. The tautomeric equilibrium constants for these proton-transfers are exceptionally sensitive to possible intramolecular interactions of the functional groups: favorable  $OH \cdots N(-N=)$  and  $NH \cdots O(C=O \text{ or } OH)$  interactions, unfavorable repulsions of lone pairs of electrons of the neighboring heteroatoms, and unfavorable repulsions of the neighboring XH groups. Therefore, the  $pK_T$  values change the sign. They vary from  $-12.9$  (**U2b**  $\rightarrow$  **U3b**) to  $+8.4$  (**U4a**  $\rightarrow$  **U5a**) for the  $N^1H \rightarrow N^3H$  conversions, and from  $-7.4$  (**U2b**  $\rightarrow$  **U4b**) to  $+7.1$  (**U3b**  $\rightarrow$  **U5b**) for the  $O^7H \rightarrow O^8H$  conversions.

Although the proton transfers from the N to C atom ( $N^1H \rightarrow C^5H$  or  $N^3H \rightarrow C^5H$ ) destroy planarity of the uracil ring and seem to be exceptionally rare, the tautomeric equilibrium constants for selected conversions are not very different from those for the favored proton-transfers from the N to O atom ( $pK_T$  10–20). Quite a different situation takes place for the proton transfers from the O to C atom ( $O^7H \rightarrow C^5H$  or  $O^8H \rightarrow C^5H$ ). Since some NH-OH tautomers have greater energy than **U7**, the  $pK_T$  values are positive or negative. They vary from  $-12.1$  (**U5b**  $\rightarrow$  **U7**) to  $+15.7$  (**U3b**  $\rightarrow$  **U7**). These observations indicate that external species (neutral molecules, ions, and even electrons), the interactions of which are energetically favorable with particular tautomers-rotamers, may change the tautomeric preferences and the tautomeric equilibrium constants for uracil<sup>3,20</sup>.

### $\pi$ -Electron Delocalization

To measure delocalization of  $\pi$ -electrons for aromatic systems, the geometry-based HOMA index (harmonic oscillator model of aromaticity) has been proposed more than thirty years ago<sup>11,12</sup> and reformulated in 1993<sup>13</sup>. However, recent applications of the reformulated HOMA index to acyclic tautomeric systems, which are less delocalized than aromatic ones, have led to the negative HOMA values for the keto and enol tautomers<sup>5,21</sup>. The negative HOMA values have been also found for the keto forms of phenols<sup>22</sup>. These observations can be explained by the fact that the reference single and double CC and CO bonds, used for calculations of the  $R_{\text{opt}}$  and  $\alpha$  values in the reformulated HOMA index, have been taken for butadiene and formic acid<sup>13</sup>, i.e. for the molecules for which  $\pi$ -electron delocalization is greater than that for keto-enol tautomers.

For this reason, we use the reference single and double CX bonds for the simplest cyclic molecules in the present work: C–C for cyclohexane, C=C for cyclohexene, C–N for piperidine, C=N for 2,3,4,5-tetrahydropyridine, C–O for cyclohexanol, and C=O for cyclohexanone. For the optimum CX bonds ( $R_{\text{opt}}$ ), we use the simplest completely delocalized molecules, benzene for CC, 1,3,5-triazine for CN, and carbonate for CO. The reference single and double CX bonds, and the optimum CX bonds ( $R_{\text{opt}}$ ) were calculated at the same level of theory {DFT(B3LYP)/6-311+G(d,p)} as the CX bonds for uracil isomers. The HOMA indices estimated in this way for all 18 possible tautomers-rotamers of uracil are given in Table IV. For each isomer, estimations were carried out for the whole tautomeric system and separately for the ring.

Generally, aromaticity is only slightly dependent on internal effects. Differences in the HOMA indices between the conformations **a** and **b** are not larger than 0.06 HOMA units.  $\pi$ -Electron delocalization seems to be mainly dependent on resonance effects, such as  $\pi$ - $\pi$ ,  $n$ - $\pi$  resonance conjugation and, additionally,  $\sigma$ - $\pi$  hyperconjugation, present only in the CH tautomers. The rare di-OH isomers (**U6a**, **U6b**, **U6c**, and **U6d**) are the most delocalized. Their HOMA indices lie between 0.86 and 0.87 for the whole tautomeric system for which the *exo*-OH groups are  $n$ - $\pi$ -conjugated with the  $\pi$ - $\pi$ -conjugated ring. For the pyrimidine ring, the HOMA indices are larger than 0.99 showing that the  $\pi$ - $\pi$  conjugation in the ring is stronger than the cross- $n$ - $\pi$  and  $\pi$ - $\pi$  conjugation between the substituents (OH) and ring.

The favored di-NH tautomer (**U1**) is moderately delocalized. Its HOMA index is equal to 0.76 for the whole tautomeric system, for which the two *endo*-N atoms with the lone electron pairs are  $n$ - $\pi$ -conjugated with the C=O

and C=C groups. For the ring, the HOMA index is slightly lower and equal to 0.70. This difference between the HOMA values suggests that the  $n$ - $\pi$  conjugation of the *endo*-N atoms with the C=C group is weaker than that with the C=O groups. A similar trend is observed for the intermolecularly H-bonded experimental structures (obtained by X-ray crystallographic and electron diffraction) reported for uracil in the literature<sup>15a,15b</sup>. The HOMA index for the ring (0.82) is slightly lower than that for the whole tautomeric system (0.85).

The rare NH-OH isomers (**U2a**, **U2b**, **U3a**, **U3b**, **U4a**, **U4b**, **U5a**, and **U5b**) are also moderately delocalized. Their HOMA indices lie between 0.65 and 0.79 for the whole tautomeric system, for which the *endo*-NH and *exo*-OH groups are cross- $n$ - $\pi$  conjugated with the C=O, C=N and/or C=C group. For the ring, the HOMA indices are not very different from those estimated for the whole system. They vary from 0.65 to 0.82. The CH isomers (**U7**, **U8a**, **U8b**, **U9a**, and **U9b**) are the least delocalized. Their HOMA indices are lower than 0.50 indicating that the  $\sigma$ - $\pi$  hyperconjugation of the ring CH<sub>2</sub> group with the C=O or C=N group is very weak, and decreases the  $\pi$ - $\pi$  and  $n$ - $\pi$  conjugation in the whole tautomeric system as well as in the ring.

A direct comparison of the decreasing HOMA order predicted for the whole tautomeric system of eighteen uracil isomers (**U6a**, **U6b**, **U6c**, **U6d**, **U4a**, **U5b**, **U3b**, **U1**, **U5a**, **U4b**, **U3a**, **U2a**, **U2b**, **U8a**, **U7**, **U8b**, **U9a**, and

TABLE IV  
HOMA indices<sup>a</sup> estimated for geometries of all 18 possible tautomers-rotamers of uracil optimized at the DFT(B3LYP)/6-311+G(d,p) level

Isomer	HOMA <sup>a</sup>		Isomer	HOMA <sup>a</sup>	
<b>U1</b>	0.755	(0.697)	<b>U6a</b>	0.872	(0.995)
<b>U2a</b>	0.691	(0.685)	<b>U6b</b>	0.870	(0.996)
<b>U2b</b>	0.653	(0.650)	<b>U6c</b>	0.865	(0.997)
<b>U3a</b>	0.741	(0.751)	<b>U6d</b>	0.863	(0.994)
<b>U3b</b>	0.758	(0.766)	<b>U7</b>	0.486	(0.349)
<b>U4a</b>	0.792	(0.817)	<b>U8a</b>	0.487	(0.394)
<b>U4b</b>	0.742	(0.766)	<b>U8b</b>	0.450	(0.346)
<b>U5a</b>	0.753	(0.768)	<b>U9a</b>	0.388	(0.283)
<b>U5b</b>	0.765	(0.784)	<b>U9b</b>	0.332	(0.222)

<sup>a</sup> For the whole tautomeric system (for the ring in parentheses).

**U9b**) with their decreasing stability order (**U1**, **U3b**, **U4a**, **U6a**, **U6b**, **U7**, **U4b**, **U3a**, **U6d**, **U6c**, **U2a**, **U5b**, **U5a**, **U2b**, **U8a**, **U9a**, **U8b**, and **U9b**) shows that aromaticity is not the main factor that dictates the tautomeric preference as it has been reported for phenols<sup>22</sup>. The favorable effects, such as intramolecular interactions between the functional groups (OH...N(-N=) and NH...O(OH or C=O)), which play an important role in the case of the rare tautomers, are also the secondary factors. The same applies for the unfavorable effects, such as repulsions of the lone electron pairs of the neighboring heteroatoms and repulsion of the neighboring XH groups. The tautomeric preference seems to be mainly dependent on stability of the functional groups. Solely this factor may explain why the diamide form **U1** is favored in the tautomeric mixture of uracil, and why the **U7** form shows greater energetic stability than the other CH tautomers (**U8a**, **U9a**, **U8b**, and **U9b**).

## CONCLUSIONS

DFT calculations performed for all possible eighteen tautomers-rotamers of uracil confirms that the tautomeric mixture contains mainly (100%) the di-NH form (**U1**). However, the NH-OH (**U3b** and **U4a**), di-OH (**U6a** and **U6b**) and CH-NH (**U7**) forms cannot be neglected to well understand some substitution, oxidation or degradation reactions. Since the decreasing dipole moment and HOMA orders are not parallel to the decreasing stability order of the eighteen isomers, polarity and aromaticity cannot be considered as the main factors that influence the tautomeric preferences for uracil as it has been suggested for model compounds (hydroxypyridines). Stability of functional groups seems to favor **U1**.

## REFERENCES

1. a) Topal M. D., Fresco J. R.: *Nature* **1976**, *263*, 285; b) Harris V. H., Smith C. L., Cummins W. J., Hamilton A. L., Adams H., Dickman M., Hornby D. P., Williams D. M.: *J. Mol. Biol.* **2003**, *326*, 1389; c) Hagen L., Peña-Díaz J., Kavli B., Otterlei M., Slupphaug G., Krokan H. E.: *Exp. Cell Res.* **2006**, *312*, 2666; d) Priet S., Gros N., Navarro J.-M., Boretto J., Canard B., Quérat G., Sire J.: *Mol. Cell* **2005**, *17*, 479; e) Bjelland S., Seeberg E.: *Mutat. Res.* **2003**, *531*, 37.
2. a) Jalbout A. F., Trzaskowski B., Xia Y., Li Y., Hu X., Li H., El-Nahas A., Adamowicz L.: *Chem. Phys.* **2007**, *332*, 152; and references therein; b) Kryachko E., Nguyen M. T., Zeegers-Huyskens T.: *J. Phys. Chem. A* **2001**, *105*, 1288; c) Tian S. X., Zhang C. F., Zhang Z. J., Chen X. J., Xu K. Z.: *Chem. Phys.* **1999**, *242*, 217.
3. Bachorz R. A., Rak J., Gutowski M.: *Phys. Chem. Chem. Phys.* **2005**, *7*, 2116.
4. Cyrański M. K., Gilski M., Jaskólski M., Krygowski T. M.: *J. Org. Chem.* **2003**, *68*, 8607.

5. Raczyńska E. D., Kosińska W., Ośmiałowski B., Gawinecki R.: *Chem. Rev.* **2005**, *105*, 3561; and references therein.
6. Parr R. G., Yang W.: *Density Functional Theory of Atoms and Molecular Orbital Theory*. Oxford University Press, New York 1989.
7. a) Becke A. D.: *J. Chem. Phys.* **1993**, *98*, 5648; b) Lee C., Yang W., Parr R. G.: *Phys. Rev. B* **1988**, *37*, 785.
8. a) Fiedler P., Böhm S., Kulhánek J., Exner O.: *Org. Biomol. Chem.* **2006**, *4*, 2003; b) Exner O., Böhm S.: *Collect. Czech. Chem. Commun.* **2006**, *71*, 1239; c) Böhm S., Fiedler P., Exner O.: *New J. Chem.* **2004**, *28*, 67.
9. a) Raczyńska E. D., Rudka T., Darowska M., Dąbkowska I., Gal J.-F., Maria P.-C.: *J. Phys. Org. Chem.* **2005**, *18*, 856; b) Raczyńska E. D., Krygowski T. M., Zachara J. E., Ośmiałowski B., Gawinecki R.: *J. Phys. Org. Chem.* **2005**, *18*, 892; c) Raczyńska E. D., Duczmal K., Darowska M.: *Vib. Spectrosc.* **2005**, *39*, 37; d) Raczyńska E. D., Duczmal K., Hallmann M.: *Pol. J. Chem.* **2007**, *81*, 1655.
10. a) Hehre W. J., Radom L., Schleyer P. v. R., Pople J. A.: *Ab initio Molecular Theory*. Wiley, New York 1986; b) Jensen F.: *Introduction to Computational Chemistry*. John Wiley & Sons, New York 1999.
11. Kruszewski J., Krygowski T. M.: *Tetrahedron Lett.* **1972**, 3839.
12. Krygowski T. M., Kruszewski J.: *Bull. Pol. Acad. Sci., Chem.* **1973**, *21*, 409.
13. Krygowski T. M.: *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 70.
14. Firsich M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Montgomery J. A., Jr., Vreven T., Kudin K. N., Burant J. C., Millam J. M., Iyengar S. S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson R., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J. E., Hratchian H. P., Cross J. B., Bakken V., Adamo C., Jaramillo R., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., Pomelli C., Ochterski J. W., Ayala P. Y., Morokuma K., Voth G. A., Salvador P., Dannenberg J. J., Zakrzewski V. G., Dapprich S., Daniels A. D., Strain M. C., Farkas O., Malick D. K., Rabuck A. D., Raghavachari K., Foresman J. B., Ortiz J. V., Cui Q., Baboul A. G., Clifford S., Cioslowski J., Stefanov B. B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R. L., Fox D. J., Keith T., Al-Laham M. A., Peng C. Y., Nanayakkara A., Challacombe M., Gill P. M. W., Johnson B., Chen W., Wong M. W., Gonzalez C., Pople J. A.: *Gaussian03*, Revision E.01. Gaussian, Inc., Wallingford, CT 2004.
15. a) Voet D., Rich A.: *Prog. Nucleic Acid Res. Mol. Biol.* **1970**, *10*, 247; b) Ferenczy G., Harsanyi L., Rozondai B., Hargittai I.: *J. Mol. Struct.* **1986**, *140*, 71; c) Ryterjans H., Kaun E., Hall W. E., Limbach H. H.: *Nucleic Acids Res.* **1982**, *10*, 7027; d) Beak P., White J. M.: *J. Am. Chem. Soc.* **1982**, *104*, 7073; e) Fujii M., Tamura T., Mikami N., Ito M.: *Chem. Phys. Lett.* **1986**, *126*, 583; f) Tsuchiya Y., Fujii M., Ito M.: *J. Phys. Chem.* **1988**, *92*, 1760; g) Szczesniak M., Nowak M. J., Szczepaniak K., Person W. B., Shugar D.: *J. Am. Chem. Soc.* **1983**, *105*, 5969; h) Padva A., O'Donnell T. J., LeBreton P. R.: *Chem. Phys. Lett.* **1976**, *41*, 278.
16. a) Exner O.: *Advances in Linear Free Energy Relationships*, Chap. 1. Plenum Press, London, New York 1972; b) Lefler J. E., Grunwald E.: *Rates and Equilibria of Organic Reactions*, Chap. 6. John Wiley and Sons, Inc., New York, London 1963.
17. a) Basch H., Garmer D. R., Jasien P. G., Krauss M., Stevens W.: *J. Chem. Phys.* **1989**, *163*, 514; b) Brown R. D., Godfrey P. D., McNaughton D., Pierlot A. P.: *J. Am. Chem. Soc.*

- 1988**, 110, 2329; c) Kulakowska I., Geller M., Lesyng B., Wierzchowski K. L., Bolewska K.: *Biochim. Biophys. Acta* **1975**, 407, 420.
18. Elguero J., Marzin C., Katritzky A. R., Linda P.: *The Tautomerism of Heterocycles*, p. 84. Academic Press, New York 1975; and references therein.
19. a) Raczyńska E. D.: *Pol. J. Chem.* **1999**, 73, 1863; b) Raczyńska E. D.: *Pol. J. Chem.* **2000**, 74, 1283.
20. a) Hobza P., Sponer J.: *Chem Rev.* **1999**, 99, 3247; b) Rejnek J., Hanus M., Labelac M., Ryjacek F., Hobza P.: *Phys. Chem. Chem. Phys.* **2005**, 7, 2006; c) Kabelac M., Hobza P.: *J. Phys. Chem. B* **2006**, 110, 14515; d) Hu X. B., Li H. R., Liang W. C., Han S. J.: *J. Phys. Chem. B* **2005**, 109, 5935; e) Hu X. B., Li H. R., Liang W. C., Han S. J.: *J. Phys. Chem. B* **2004**, 108, 12999; f) Zhang L., Li H., Hu X., Jalbout A. F.: *Chem. Phys.* **2007**, 337, 110.
21. a) Raczyńska E. D.: *Pol. J. Chem.* **2005**, 79, 749; b) Raczyńska E. D.: *Pol. J. Chem.* **2005**, 79, 1003.
22. Ośmiałowski B., Raczyńska E. D., Krygowski T. M.: *J. Org. Chem.* **2006**, 71, 3727.