

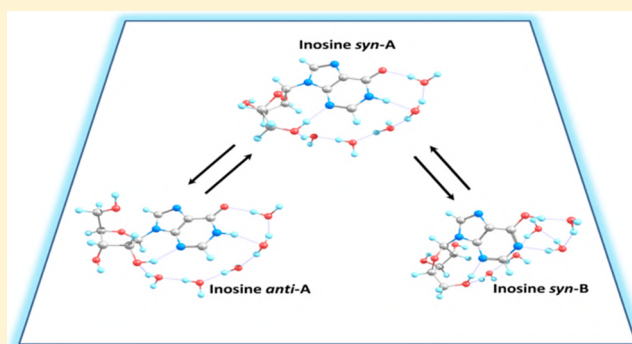
# Tautomerism of Inosine in Water: Is It Possible?

Nadezhda Markova and Venelin Enchev\*

Institute of Organic Chemistry with Center of Phytochemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

## Supporting Information

**ABSTRACT:** Syn- and anti-conformers of four tautomer structures of inosine were studied in the gas phase and in solvent water to investigate the possibility of hydrogen bonding and tautomeric conversion. It was found that in the gas phase and in water solution the most stable is the syn-conformer of the 6-keto tautomer followed by its anti-conformer and syn-conformer of the 6-enol form. In the gas phase, the percent content of syn- and anti-conformers is 83.77 and 12.86%, respectively, whereas syn-enol tautomer is calculated to be 2.54%. However, in water solution the syn-conformer of the keto tautomer is 99.9%. The water-assisted proton transfer process in inosine was investigated using ab initio MP2 and SCS-MP2 quantum chemical approaches. Solute–solvent clusters containing an inosine molecule and five water molecules embedded in the “bulk” solvent treated as polarizable continuum (the CPCM/MP2/6-31+G(d,p) level of theory) were explored. The energy barrier of the water-assisted proton transfer reaction in inosine is found to be 12.9 kcal mol<sup>−1</sup> and the rate constant ( $k = 6.68 \times 10^1 \text{ s}^{-1}$ ) is sufficiently large to generate the 6-enol tautomer. From the reaction profile of the proton transfer we can conclude that the process is realized through the asynchronous concerted mechanism.



## 1. INTRODUCTION

Inosine, 9-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)-oxolan-2-yl]-3H-purin-6-one, is a purine nucleoside that has hypoxanthine linked by N9 nitrogen to C1 carbon ( $\beta$ -N<sub>9</sub>-glycosidic bond) of ribose. It is an intermediate in the degradation of purines and purine nucleosides to uric acid and in the pathways of purine salvage. Inosine is commonly found in tRNAs and is essential for proper translation of the genetic code in wobble base pairs.

Tautomerism of inosine is considered in several papers.<sup>1–5</sup> Wolfenden<sup>2</sup> has suggested that the enol tautomer of inosine, if it is stable, could form base pairs of the Watson–Crick type with uridine and account for the occurrence of inosine–uridine pairs between the codon and anticodon bases without disrupting the normal relationship between the glycosidic bonds. The appearance of rare tautomers can lead to spontaneous mutations owing to the base mispairing. Topal and Fresco<sup>6</sup> proposed an alternative base-pairing scheme for adenosine–inosine as a more suitable codon–anticodon interaction.

On the basis of the infrared spectra of inosine, Miles<sup>1</sup> concluded that the 6-keto form predominates. By means of Raman spectroscopy, Medeiros and Thomas<sup>3</sup> checked the possibility of 6-keto and 6-enol tautomers of inosine to coexist in aqueous solution. The spectra of aqueous solutions of inosine, in the region 2000–200 cm<sup>−1</sup>, indicate that inosine exists predominantly in the keto form.

In the solid state, inosine crystallized in three different crystal forms<sup>7</sup>—two monoclinic<sup>8–10</sup> and one orthorhombic

forms.<sup>10,11</sup> The crystals with monoclinic symmetry<sup>8</sup> show two molecules in the unit cell and a special arrangement of intermolecular hydrogen bonds. In the orthorhombic form,<sup>11</sup> the two crystallographically independent molecules are essentially identical with regard to bond lengths, bond angles, and conformational parameters. The location in the crystal of a proton on N1 confirms the tautomeric form as amino-oxo tautomer A (Figure 1) or the so called 6-keto form for both molecules.

The crystal structure of inosine dihydrate<sup>9</sup> has also been reported. In this case, the monoclinic crystals show two nucleoside molecules (amino-oxo tautomer A) and four water molecules per asymmetric unit. Chiarella et al.<sup>12</sup> have studied the crystal nucleation of inosine. According to proton NMR spectra, the initial nucleation of a metastable polymorph at temperatures above 10 °C has been directed by dimeric self-association. At the lower temperatures, at 10 °C and below, the dihydrate has been the most stable phase and has become the first to appear with no evidence for the existence of either of the polymorphs.

The low frequency vibrations (250–400 cm<sup>−1</sup> region) of the polycrystalline powder of inosine at room temperature were measured and the vibrations associated with the ribose ring were determined.<sup>13</sup> The spectrum of inosine has been compared with the spectrum of hypoxanthine. The ribose

Received: November 22, 2018

Revised: December 28, 2018

Published: January 3, 2019



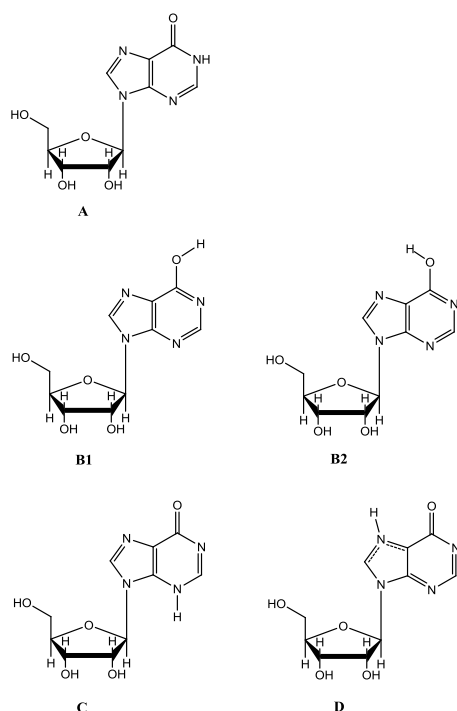


Figure 1. Inosine tautomers and rotamers.

vibrations of inosine are not clearly distinguished because the hypoxanthine has several bands in the same spectral region as those of the ribose.

Many theoretical papers have been devoted to the conformational isomerism of DNA and RNA structural components,<sup>14–19</sup> but few studies appear on inosine. The prototropic tautomerism and basic molecular principles of hypoxanthine mutagenicity were investigated to understand the elementary molecular mechanisms of mutagenic action of hypoxanthine as a product of adenine deamination in DNA.<sup>15</sup> Five tautomers of inosine were determined and optimized at the MP2 and B3LYP levels of theory and a comprehensive conformational analysis was carried out on the most stable keto tautomer.<sup>20</sup> It has been found that the least stable tautomer (the H-atom is positioned to nitrogen from the five-membered ring of the purine moiety) in the isolated state is the most favored in a polarizable environment with water because of its dipole moment among all five tautomers. Although tautomerism of nucleobases and nucleosides has been studied in the gas phase, aprotic solvents or excited state, the conclusions are less relevant for DNA and RNA structural components. Under these conditions tautomeric equilibria are significantly altered and the relative proportion of major and minor tautomeric forms changed, even for the canonical nucleobases.<sup>21</sup> To study tautomerism in aqueous solutions is more difficult because tautomerization is mediated by water molecules. This process leads to fast rates of the tautomeric equilibria and low abundance of minor tautomeric species.<sup>22</sup>

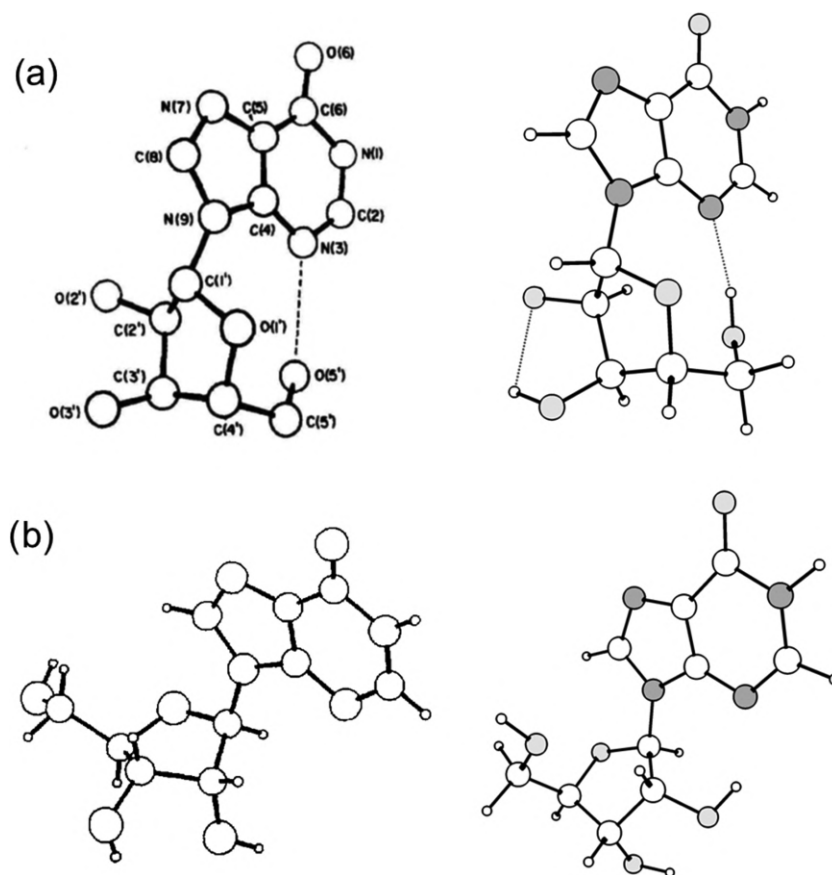


Figure 2. Structure of the tautomeric form A of inosine. (a) The syn-conformer and (b) anti-conformer. The X-ray diffraction structure of the syn-form, adapted by ref 11 and the anti-conformer, adapted by ref 8 are presented in the left. MP2 optimized structures are shown in the right. The dotted lines depict the intramolecular H-bonds.

**Table 1.** Relative Free Energies,  $\Delta G_{298}$  (kcal mol<sup>-1</sup>), and the Percent Content ( $T\%$ ) for the Syn- and Anti-Conformers (Figure 2) of Inosine Tautomers (Figure 1) in the Gas Phase and in Water Solution, Calculated at MP2/6-31+G(d,p) and SCS-MP2/6-31+G(d,p) Levels<sup>a</sup>

molecule	inosine		inosine (CPCM)		inosine + 5H <sub>2</sub> O (CPCM)	
	$\Delta G_{298}$	$T\%$	$\Delta G_{298}$	$T\%$	$\Delta G_{298}$	$T\%$
MP2						
Is-A-syn	0.00	83.77	0.00	99.99	0.00	99.96
Is-A-anti	1.11	12.86	5.75	$6.1 \times 10^{-3}$	4.27	0.04
Is-B1-syn	2.07	2.54	6.37	$2.1 \times 10^{-3}$	7.17	$1.3 \times 10^{-4}$
Is-B1-anti	3.24	0.35	11.65	$1.0 \times 10^{-7}$	10.77	$9.9 \times 10^{-8}$
Is-B2-syn	3.19	0.38	9.39	$1.3 \times 10^{-5}$	12.96	$2.3 \times 10^{-9}$
Is-B2-anti	4.08	0.09	12.83	$3.9 \times 10^{-8}$	12.64	$4.2 \times 10^{-9}$
Is-C-syn	23.01	$1.1 \times 10^{-15}$	18.12	$5.2 \times 10^{-12}$	15.62	$1.5 \times 10^{-11}$
Is-C-anti	23.00	$1.2 \times 10^{-15}$	18.51	$2.7 \times 10^{-12}$	12.04	$1.3 \times 10^{-8}$
Is-D-syn	20.46	$8.4 \times 10^{-14}$	14.47	$2.5 \times 10^{-9}$	11.14	$7.2 \times 10^{-8}$
Is-D-anti	19.27	$6.3 \times 10^{-13}$	15.58	$3.8 \times 10^{-10}$	11.85	$1.9 \times 10^{-8}$
TS(A → B1)-syn	36.08				9.71	
TS(syn-anti)	4.89		8.60			
SCS-MP2						
Is-A-syn	0.00	87.90	0.00	99.99	0.00	99.91
Is-A-anti	1.21	11.40	5.83	$5.3 \times 10^{-3}$	3.70	0.09
Is-B1-syn	3.02	0.55	7.27	$4.6 \times 10^{-4}$	8.60	$8.6 \times 10^{-6}$
Is-B1-anti	4.22	0.07	12.60	$5.8 \times 10^{-8}$	11.64	$2.8 \times 10^{-8}$
Is-B2-syn	4.18	0.07	10.31	$2.6 \times 10^{-6}$	13.61	$6.8 \times 10^{-10}$
Is-B2-anti	5.11	0.01	13.78	$7.8 \times 10^{-9}$	13.44	$9.3 \times 10^{-10}$
Is-C-syn	22.06	$5.9 \times 10^{-15}$	17.57	$1.3 \times 10^{-11}$	15.13	$3.8 \times 10^{-11}$
Is-C-anti	21.95	$7.1 \times 10^{-15}$	18.00	$6.4 \times 10^{-12}$	12.25	$8.8 \times 10^{-9}$
Is-D-syn	21.37	$1.9 \times 10^{-14}$	15.05	$9.3 \times 10^{-11}$	12.09	$1.2 \times 10^{-8}$
Is-D-anti	20.09	$1.6 \times 10^{-13}$	16.12	$1.5 \times 10^{-10}$	12.53	$5.2 \times 10^{-9}$
TS(A → B1)-syn	38.88				12.90	
TS(syn-anti)	4.33		8.04			

<sup>a</sup>Energy barriers of rotation of the ribose ring in syn- and anti-conformers are given.

Since the inosine tautomerism in water is not studied enough, the main aim of the present study is to investigate the possibility for water-assisted proton transfer in the title compound. In the processes that occur in water solutions, the solvent influence has a considerable effect.

## 2. COMPUTATIONAL DETAILS

The medium-level MP2 calculations provide qualitatively correct results for molecular complexes of nucleic acid bases. The geometries, normal mode vibrational frequencies, and Raman activities of the inosine tautomers were computed at the second order Møller–Plesset level using the 6-31+G(d,p) basis set.<sup>23,24</sup> The conductor polarizable continuum model (CPCM) formalism<sup>25</sup> was used to simulate the solvent influence. Solute–solvent clusters containing an inosine molecule and five water molecules embedded in polarizable continuum (the CPCM/MP2/6-31+G(d,p) level of theory) were explored. The tautomer geometries were also optimized at the CPCM/MP2/6-31+G(d,p) level and solution-phase vibrational frequencies were computed at the same computational level. All optimized structures showed only positive harmonic vibrations (local energy minima). The transition state was confirmed to have one imaginary frequency. The transition structure was checked by intrinsic reaction coordinate calculations. Starting from the transition state, the reaction path was generated as the steepest descent path in mass-scaled coordinates (intrinsic reaction coordinate, IRC) using the Gonzalez–Schlegel algorithm,<sup>26</sup> employing a step size of 0.05 bohr (1 bohr corresponds to 0.53 Å).

The spin-component-scaled MP2 (SCS-MP2) method<sup>27</sup> was used for single-point energy calculations. It provides significantly improved energetics compared to standard MP2 reaching QCISD(T) accuracy. For comparison, the Gibbs energy difference between *cis*-2-hydroxypyridine and 2-pyridone is calculated to be 4.6 kJ mol<sup>-1</sup> at the SCS-MP2/aug-cc-pVTZ level,<sup>28</sup> which is closer to the experimental value, 3.2 kJ mol<sup>-1</sup>, than the value calculated at the CCSD(T)/aug-cc-pVTZ level, 5.3 kJ mol<sup>-1</sup>. SCS-MP2 has also the advantage of lower computational cost than QCISD(T) and CCSD(T).

The Gibbs free energy,  $G$ , for all structures was obtained as sum:  $G = E_t + E_{\text{corr}}$ , where  $E_t$  is the total (electronic + nuclear) energy and  $E_{\text{corr}}$ —thermal correction. For the calculation of the zero-point vibrational energies and thermal correction to total energy, the frequencies were retained unscaled. The values of Gibbs free energies ( $\Delta G$ ) and activation barriers ( $\Delta G^\ddagger$ ) were calculated at temperature 298.15 K.

The percent content of the inosine tautomeric forms and their syn- and anti-conformers,  $T\%$ , was calculated as

$$T\% = 100 \cdot p_i, \text{ where } p_i = e^{-\Delta G_i/RT} / \sum_i e^{-\Delta G_i/RT} \quad (1)$$

The rate constant of the forward ( $k_f$ ) and the reverse ( $k_r$ ) reaction was calculated by the Eyring equation

$$k_{fr} = \frac{k_B T}{h} e^{-\Delta G_{fr}^\ddagger/RT} \quad (2)$$

where  $k_B$ ,  $h$ , and  $R$  are the Boltzmann, Planck, and universal gas constants, respectively;  $\Delta G_{fr}^\ddagger$  is the Gibbs free energy of

Table 2. MP2/6-31+G(d,p) Calculated and Experimental Bond Lengths (in Å) for Anti- and Syn-Conformers of Inosine Tautomer A<sup>a</sup>

parameter	anti-A			syn-A	
bond	exptl. <sup>9</sup>	exptl. <sup>8</sup>	calculated	exptl. <sup>11</sup>	calculated
N1–C2	1.360	1.355	1.365	1.335	1.365
C2–N3	1.307	1.308	1.313	1.290	1.312
N3–C4	1.362	1.365	1.373	1.364	1.373
C4–C5	1.382	1.374	1.402	1.368	1.400
C5–C6	1.429	1.433	1.443	1.430	1.445
C6–N1	1.401	1.397	1.432	1.432	1.429
C6–O6	1.221	1.233	1.229	1.202	1.228
C5–N7	1.371	1.371	1.371	1.390	1.373
N7–C8	1.311	1.307	1.334	1.299	1.327
C8–N9	1.359	1.377	1.373	1.365	1.377
C4–N9	1.372	1.372	1.370	1.381	1.371
N9–C1'	1.383	1.477	1.467	1.460	1.445
C1'–O1'	1.460	1.417	1.411	1.426	1.416
C1'–C2'	1.525	1.530	1.542	1.513	1.528
C2'–C3'	1.517	1.525	1.526	1.534	1.530
C3'–C4'	1.526	1.522	1.523	1.539	1.528
O1'–C4'	1.452	1.459	1.452	1.443	1.462
C2'–O2'	1.417	1.420	1.418	1.410	1.426
C3'–O3'	1.418	1.413	1.423	1.430	1.427
C4'–C5'	1.513	1.508	1.513	1.497	1.515
C5'–O5'	1.437	1.428	1.431	1.417	1.422
O2'–H			0.977		0.966
O3'–H			0.971		0.971
O5'–H			0.966		0.979
N1'–H			1.014		1.014
O5'...N3				2.80	2.89
Torsional Angle					
C8–N9–C1'–C2'		–121.2	–127.3	121.9	120.8
O1'–C4'–C5'–O5'			66.5	62.7	69.9
C3'–C4'–C5'–O5'			52.3	56.7	47.5

<sup>a</sup>For numbering of the atoms see Figure 2.

activation for the reaction in the forward (f) and reverse (r) directions and  $T = 298.15$  K.

The present ab initio calculations were performed using GAMESS<sup>29,30</sup> suite of programs.

### 3. RESULTS AND DISCUSSION

**3.1. Tautomeric and Conformational Rearrangements in Inosine.** Similar to other nucleosides (guanosine, adenosine) several tautomeric forms are possible in inosine. Four tautomeric forms, three amino-oxo (A, C, and D), and one imino-hydroxy (B), are shown in Figure 1. The tautomer B has two rotamers: B1 and B2. Five structural parameters should be taken into account for a complete description of the conformation of inosine. Following the Saenger's notation<sup>31</sup> the atomic description of tautomer A is defined in Figure 2. The glycosidic torsional angle,  $\chi(C4-N9-C1'-O1')$ , determines the base position with respect to the ribose plane and by this way two conformations, anti (see Figure 2a) and syn (see Figure 2b), can be generated. The exocyclic torsional angles  $\gamma(C3'-C4'-C5'-O5')$  and  $\beta(C4'-C5'-O5'-H5')$  describe the orientation of the  $C5'-O5'$  and  $O5'-H5'$  groups with respect to the furanose ring, respectively. Tautomers A, C, and D, and rotamers B1 and B2 have two conformers.

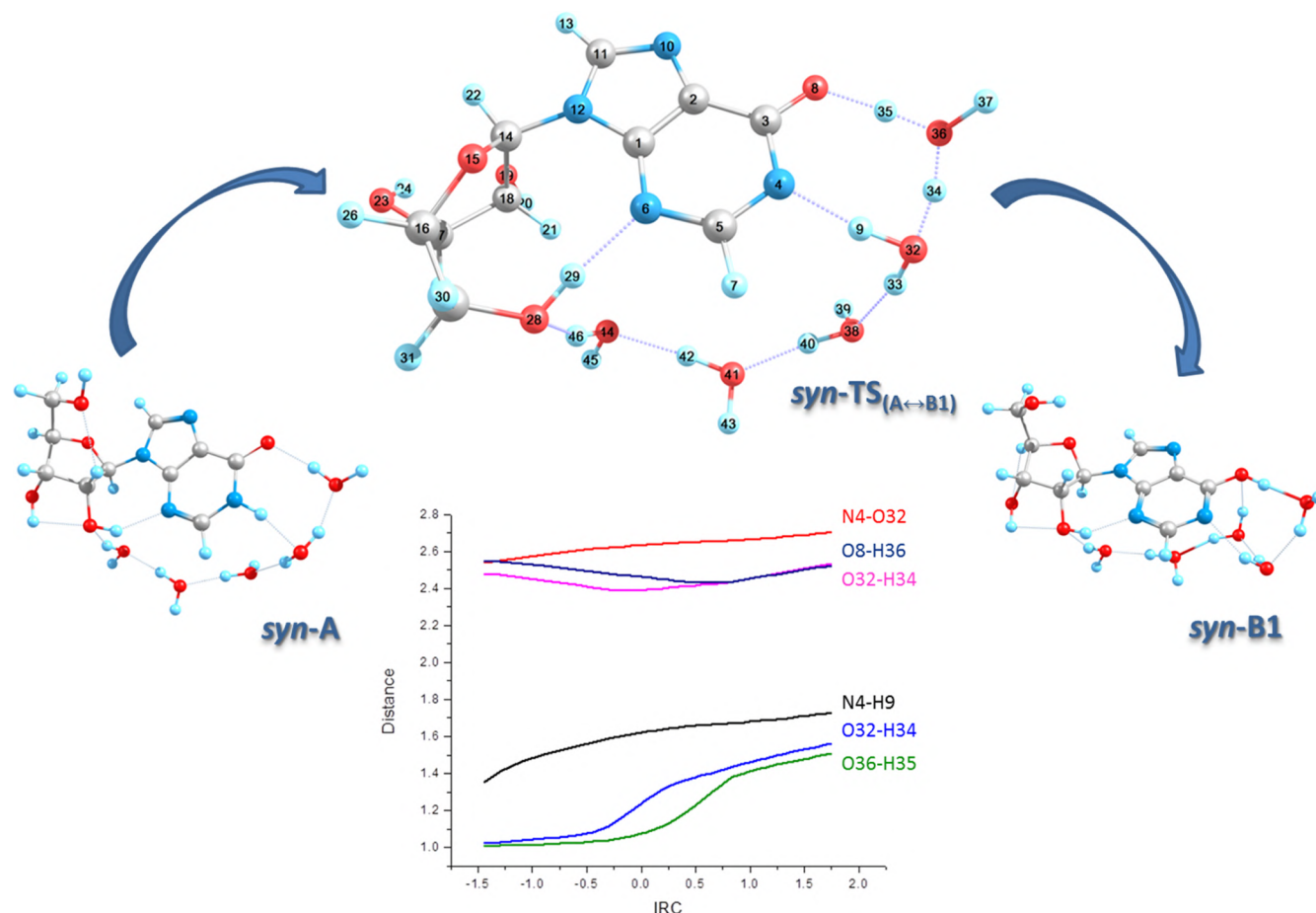
Since inosine can have a different number of conformations, our starting geometries were selected based on the conformers observed in crystal structures.<sup>8,11</sup>

The relative Gibbs free energies of syn- and anti-conformers of the inosine tautomers computed at the MP2/6-31+G(d,p) level are presented in Table 1. Additional single-point calculations at the SCS-MP2/6-31+G(d,p) level were carried out.

The computed energies of the inosine tautomers in the gas phase reveal that the syn-conformer of tautomer A is most stable followed by its anti-conformer as the energy difference between both conformers is  $1.11 \text{ kcal mol}^{-1}$  (Table 1). Tautomer syn-B1 is higher in energy by  $2.07 \text{ kcal mol}^{-1}$ . The energy difference between rotamers syn-B1 and syn-B2 is  $1.12 \text{ kcal mol}^{-1}$ . The quantity of these species, according to their calculated relative stabilities (eq 1) amount to 83.7% for syn-A, 12.86% for anti-A, 2.54% for syn-B1, 0.38% for syn-B2, 0.35% for anti-B1, and 0.09% for anti-B2. The calculated energy barrier of the rotation of the furanose ring around the N9–C1' glycosidic bond is low, that is,  $4.89 \text{ kcal mol}^{-1}$ . The probable parallel reaction, intramolecular proton transfer syn-A  $\rightarrow$  syn-B1, could not occur because its calculated energy barrier is too high, that is,  $36.08 \text{ kcal mol}^{-1}$ .

The calculations at the SCS-MP2/6-31+G(d,p) level show similar qualitative results. However, the energy difference between conformers syn-A and anti-A increases to  $1.21 \text{ kcal mol}^{-1}$ , whereas the energy barrier of the rotation around the glycosidic bond is practically the same, that is,  $4.33 \text{ kcal mol}^{-1}$ . The energy gap between the tautomers syn-A and syn-B1





**Figure 3.** Structures of the five-hydrated tautomeric forms *syn-A* and *syn-B1* of inosine and the respective transition state calculated at the CPCM/MP2/6-31+G(d,p) level, and the distances between the two selected atoms (in Å) along the IRC (in  $\text{amu}^{1/2} \text{ bohr}$ ).

increases to  $3.02 \text{ kcal mol}^{-1}$  and the intramolecular proton transfer *syn-A*  $\rightarrow$  *syn-B1* also increases to  $38.88 \text{ kcal mol}^{-1}$  (Table 1). In this way the quantity of *syn-A*, anti-*A*, and *syn-B1* in the gas phase is found to be 87.9, 11.4, and 0.54%, respectively.

*Syn-* and anti-conformers of the amino-oxo tautomer *A* have been isolated in the crystal form from water solution.<sup>8,11</sup> The calculated and available experimental bond distances for both conformers of *A* are shown in Table 2. For the anti-conformer, the MP2/6-31+G(d,p) calculations predict C1–C2 and C2–C3 bonds to be longer by 0.020 and 0.014 Å, respectively than the experimental ones in the hypoxanthine moiety. The calculated C3=O8 bond is longer by 0.008 Å than the experimentally observed one. The largest deviation is observed for the C3–N4 bond in the six-membered ring and N10–C11 one in the five-membered ring where the calculated distances are longer by 0.031 and 0.023 Å, respectively. Similar results are obtained for the *syn*-conformer. In the optimized structure, the C3=O8 bond is shorter by 0.026 Å and C1–C2 and C2–C3 bonds are longer by 0.032 and 0.015 Å than the experimental ones. Larger deviations are observed for carbon–nitrogen bonds. C5–N4, C5–N6, and C11–N10 are longer by 0.032, 0.022, and 0.028 Å, respectively, whereas C2–N10, C1–N12, and C11–N12 are shorter by 0.017, 0.010, and 0.015 Å, respectively, than the experimentally observed one.

The relative Gibbs free energies of the tautomers of inosine in water solution (CPCM) are presented in Table 1. The

energy difference between both conformers of tautomer *A* increases to  $5.75 \text{ kcal mol}^{-1}$  and the difference between the *syn*-conformers of tautomers *A* and *B1* is calculated to be  $6.37 \text{ kcal mol}^{-1}$ . The quantities of these three species amounts to 99.99% *syn-A*,  $6.1 \times 10^{-3}\%$  anti-*A*, and  $2.1 \times 10^{-3}\%$  *syn-B1* (Table 1). The calculated energy barrier of the rotation *syn-A*  $\rightarrow$  anti-*A* increases to  $8.60 \text{ kcal mol}^{-1}$ , i.e., it is 57% higher in comparison to the gas phase.

The results are practically the same when we take into account the improved CPCM/SCS-MP2/6-31+G(d,p) energetics compared to CPCM/MP2/6-31+G(d,p) ones—99.99% *syn-A*,  $5.3 \times 10^{-3}\%$  anti-*A*, and  $4.6 \times 10^{-4}\%$  *syn-B1*. The predicted barrier of rotation *syn-A*  $\rightarrow$  anti-*A* is calculated to be  $8.04 \text{ kcal mol}^{-1}$ , which is almost 2 times higher than those in the gas phase calculated at the same computational level of theory.

It is known that in aqueous solution, the solvent can form intermolecular hydrogen bonds with solute (short-range interactions) and can also act as a catalyst of proton transfer reactions. For simulation of these short-range interactions in aqueous medium, complexes containing five water molecules around the all considered tautomers, rotamers, and conformers of inosine were used. To take into account the impact of the long-range interactions of the investigated complexes, their geometry optimizations were performed at the CPCM/MP2/6-31+G(d,p) level, choosing water as surrounding media. The nature of the stationary points or TS as minima or maxima on

**Table 3.** CPCM/MP2/6-31+G(d,p) Calculated Bond Lengths (in Å) for Water Complexes of Inosine Tautomers A and B1<sup>a</sup>

bonds	syn-A + 5w	syn-B1 + 5w	syn-TS <sub>(A→B1)</sub> + 5w
C1–C2	1.399	1.405	1.400
C2–C3	1.431	1.405	1.426
C3–N4	1.396	1.338	1.371
C3–O8	1.251	1.334	1.287
N4–C5	1.360	1.355	1.351
C5–N6	1.318	1.335	1.333
C6–C1	1.365	1.349	1.351
C1–N12	1.375	1.379	1.380
C2–N10	1.376	1.378	1.381
N10–C11	1.330	1.329	1.329
N12–C11	1.373	1.374	1.374
C12–N14	1.444	1.447	1.445
C14–O15	1.417	1.416	1.418
O15–C16	1.460	1.466	1.463
C16–C17	1.525	1.528	1.528
C17–C18	1.533	1.525	1.531
C18–O19	1.423	1.420	1.423
C17–O23	1.428	1.428	1.429
C14–C18	1.525	1.521	1.525
C16–C27	1.513	1.528	1.528
C27–O28	1.428	1.432	1.429
O19–H20	0.969	0.970	0.969
O23–H24	0.973	0.974	0.973
O28–H29	0.989	0.990	0.997
N4–H9	1.039	1.949	1.620
O8–H35		1.005	1.383
N4–O32			2.633
O32–O36			2.391
O32–H34			1.237
O36–H34			1.170
O8–O36			2.461
O36–H35			1.075

<sup>a</sup>For numbering of the atoms see Figure 3.

the potential energy surface was checked again by the absence or the presence, respectively, of imaginary frequency.

Using this computational model at the CPCM/MP2/6-31+G(d,p) level it is found that the energy difference between both conformers (syn-A and anti-A) is 4.27 kcal mol<sup>−1</sup> and that between tautomers syn-A and syn-B1 is calculated to be 7.17 kcal mol<sup>−1</sup>, but the tautomeric conversion is characterized with a low energy barrier for forward (9.71 kcal mol<sup>−1</sup>) and reverse (2.54 kcal mol<sup>−1</sup>) reactions.

According to CPCM/SCS-MP2/6-31+G(d,p) calculations, the energy difference between the syn- and anti-conformers of tautomer A is 3.70 kcal mol<sup>−1</sup>, whereas that between syn-tautomers A and B1 is found to be 8.60 kcal mol<sup>−1</sup>. The percent content of the three species amount to 99.91% syn-A, 0.09% anti-A, and  $8.6 \times 10^{-6}\%$  syn-B1 (Table 1). The energy barrier of water-assisted intramolecular proton transfer syn-A → syn-B1 is calculated to be 12.9 kcal mol<sup>−1</sup>.

The reaction syn-A → syn-B1 starts as the proton H9 rapidly moves from nitrogen atom N4 to the oxygen atom O32 and the N4–O32 distance begins to elongate (Figure 3 and Table 3), i.e., the water molecule situated around the N4 atom of inosine moves away from nucleoside. At the same time O32...O36 and O8...O36 distances slowly decrease. The equilibrium distance O32...O36 shortens from 2.718 Å from its

equilibrium value to 2.391 Å in the transition state, and after that slowly elongates. The transition state is formed when the distances N4–H9, O32–H34, and O36–H35 become 1.620, 1.237, and 1.075 Å, respectively. The carbonyl bond C3=O8 is elongated from 1.251 to 1.287 Å and the single C3–N4 bond is shortened from 1.396 to 1.371 Å. Parallel to the shortening of O32...O36, protons H34 and H35 start a rapid transfer to oxygen atoms O36 and O8, respectively. In the beginning of the reaction the distance O8...O36 decreases from its equilibrium value 2.770–2.481 Å. When the O8–O36 distance becomes 2.481 Å, the protons H34 and H35 begin to transfer slowly, whereas the bonds O8–H35 and O36–H34 are formed.

The calculated forward rate constant of tautomerization,  $k_f = 6.68 \times 10^1 \text{ s}^{-1}$  (eq 2), is enough to generate a concentration of the tautomeric form syn-B1. The rate constant of reverse reaction is  $k_r = 4.98 \times 10^7 \text{ s}^{-1}$  and the time necessary to reach the equilibrium concentration between the tautomeric forms syn-A and syn-B1 was calculated by the formula:  $\tau_{99.9\%} = \ln 10^3 / (k_f + k_r)$  and it is  $1.38 \times 10^{-7} \text{ s}$ .

**3.2. Infrared and Raman Spectral Analysis.** Spectroscopic methods for the study of tautomerism of DNA and RNA structural components in aqueous solutions have been of limited use. Vibrational IR and Raman spectroscopies are sensitive to identifying tautomeric forms of nucleobases and nucleosides in aqueous solutions.

Due to the solvent absorption in the region 1800–1500 cm<sup>−1</sup>, the possibility to use infrared spectroscopy is limited. The Raman spectroscopy of aqueous inosine in the region 2000–200 cm<sup>−1</sup> is more appropriate<sup>32</sup> because liquid water gives rise to only very weak Raman scattering over this region.

The new developments of spectroscopic methods, such as two-dimensional IR<sup>33</sup> and variable temperature NMR spectroscopies, combined with quantum chemical calculations allow comprehensive characterization of tautomeric equilibria in solution.<sup>34</sup> The available vibrational IR and Raman spectra indicate that in neutral aqueous solution inosine exists predominantly in the keto form.<sup>3</sup> Conventional Raman spectroscopy has proven to be a useful tool for probing the nucleic acid structure. Since the sugar rings exhibit only relatively weak Raman signals, Bell et al.<sup>35</sup> have focused mainly on the intense characteristic bands of the nucleic acid bases.

We have calculated IR frequencies, intensities, and Raman activities of inosine conformers (syn- and anti-) for tautomers A and B in the gas phase and water. We considered the effect of aqueous hydration on IR and Raman spectra of inosine using an approach in which the cluster including water molecules and inosine was embedded in polar medium (water) and the CPCM/MP2/6-31+G(d,p) calculations were performed. The cluster including five water molecules forming a chain between the C=O group from the hypoxanthine moiety and the OH-group from the ribose ring are attached to the tautomers A and B of the inosine syn- and anti-conformers. According to our results there is a good agreement between theoretical and experimental data. Table 4 exhibits the selected infrared frequencies and intensities as well as Raman activities of inosine.

In the range from 1800 to 1500 cm<sup>−1</sup>, strong bands are located because of the C=O stretching, skeletal stretching, and NH bending vibrations of the organic base residues.<sup>36</sup> The most intensive band for tautomer A for both conformers in the gas phase was calculated to be at 1709 cm<sup>−1</sup> (anti-) and 1710 cm<sup>−1</sup> (syn-). This corresponds well to the experimentally

**Table 4.** MP2/6-31+G(d,p) Calculated Frequencies, Scaled by 0.945, are in  $\text{cm}^{-1}$ , Intensities,  $I$  ( $\text{km mol}^{-1}$ ) and Raman activities,  $A$  ( $\text{\AA}^4 \text{amu}^{-1}$ ) for Syn- and Anti-Conformers of Inosine Tautomers A and B (Figure 3)<sup>a</sup>

anti		experimental		syn		assignment
isolatedfreq. ( <i>I/A</i> )	CPCM/clusterfreq. ( <i>I</i> )	freq. ( <i>I</i> ) <sup>40</sup>	freq. ( <i>A</i> ) <sup>3</sup>	isolatedfreq. ( <i>I/A</i> )	CPCM/clusterfreq. ( <i>I</i> )	
Tautomer A						
3440 (98/149)	3144 (1247)	3432		3439 (79/130)	2984 (1734)	$\nu_{\text{N3-H2}}$
1709 (700/127)	1651 (1064)	1702	1677 (3)	1710 (704/126)	1659 (1011)	$\nu_{\text{C=O}} + \nu_{\text{C3-C2}} + \delta_{\text{N4-H9}} + \delta_{\text{C3-N4}} + \gamma_{\text{H2O}}$
	1570 (281)	1580	1581 (3)			$\delta_{\text{NH}}$
1554 (110/95)	1567 (132)	1544	1549 (9)	1554 (121/99)	1579 (156)	$\delta_{\text{C5-H7}} + \nu_{\text{C5-N4}} + \nu_{\text{C3-N4}}$
1406 (0.1/22)		1380		1397 (58/62)		$\delta_{\text{N4-H9}}$
1385 (44/30)	1488 (3)	1373	1507 (10)	1415 (45/64)	1480 (10)	$\delta_{\text{NH}} + \nu_{\text{C11-N12}}$
1375 (2/279)				1390 (28/23)		$\delta_{\text{NH}} + \delta_{\text{C-H}} + \nu_{\text{C5-N4}}$
1360 (49/4)	1377 (207)	1349	1390 (2)	1361 (51/23)	1377 (184)	$\nu_{\text{C2-N10}} + \delta_{\text{C=O}} + \delta_{\text{C3-C2}} + \nu_{\text{C3-C2}}$
	1346 (66)		1336		1352 (11)	$\delta_{\text{C-H}}$
1113 (41/5)		1095		1117 (40/3)		$\delta_{\text{N4-H9}} + \nu_{\text{C3-N4}} + \nu_{\text{C5-N4}}$
1055 (28/1)		1066		1071 (19/4)		$\nu_{\text{C3-N4}}$
Tautomer B						
3575 (132)	2924 (2862)			3574 (135)	2922 (2696)	$\nu_{\text{O9-H8}}$
1537 (127)	1525 (226)			1544 (168)	1597 (377)	$\delta_{\text{C=N}} + \nu_{\text{C6-N1}} + \nu_{\text{C1-C2}}$
	1385 (449)				1381 (223)	$\nu_{\text{C3=N4}} + \nu_{\text{C3-O8}}$
1430 (163)				1436 (128)		$\nu_{\text{C3=N4}} + \delta_{\text{C5-H7}}$
1257 (22)	1423 (273)			1260 (13)	1167 (134)	$\delta_{\text{O8-H9}}$

<sup>a</sup>Available experimental data are also presented.

found band at  $1702 \text{ cm}^{-1}$  for  $\text{C=O}$  stretching in IR spectra of inosine in KBr. When the water surrounding was taken into account in the theoretical IR and Raman spectra, the band near  $1680 \text{ cm}^{-1}$  is predominantly due to the carbonyl stretching vibration<sup>1</sup> and identifies the keto structure. The theoretically predicted line for this vibration of syn- and anti-conformers was found to be at  $1659$  and  $1651 \text{ cm}^{-1}$ , respectively and they correspond to the weak bands at  $1677 \text{ cm}^{-1}$  in the Raman spectra.

Solvent absorption (water) makes it impossible to record the spectra over  $3400 \text{ cm}^{-1}$  and in this way the N–H (keto form A) and O–H (enol form B) stretching frequencies cannot be registered.<sup>4</sup> The theoretically predicted band  $\nu_{\text{N3-H2}}$  for syn- and anti-conformers of A was found to be  $3440$  and  $3439 \text{ cm}^{-1}$ , which is close to the experimentally observed band at  $3432 \text{ cm}^{-1}$ . The line corresponds to the O–H stretching frequency of the two conformers of imino-hydroxy tautomer B that was calculated to be at  $3575 \text{ cm}^{-1}$  (anti-) and  $3574 \text{ cm}^{-1}$  (syn-). When IR spectra were simulated in water solution, the bands for N–H and O–H stretching frequencies were observed in the region around  $3000 \text{ cm}^{-1}$  for the two tautomers in syn- and anti-conformations.

The experimentally found  $1579$  and  $1509 \text{ cm}^{-1}$  bands from infrared absorption spectra in  $\text{H}_2\text{O}$  of inosine at  $\text{pH} = 6.3$  are due to the characteristic vibrations of the purine ring and lines at or near these frequencies are present also in the Raman spectra of adenine and guanine derivatives.<sup>32</sup> On the basis of a detailed study of the infrared spectra in water of inosine, these lines may be assigned to the C–N stretching and N–H bending frequencies. According to the theoretically calculated frequencies in water solution, two near lines (at  $1570$  and  $1567 \text{ cm}^{-1}$ ) were observed in the spectrum of the anti-conformer of A, the first of them were absent from the spectrum of the syn-conformer of A and may be ascribed to the N–H bending and mixture of frequencies assigned to the water molecules. The line at  $1567 \text{ cm}^{-1}$  was found in the syn-conformer of A at  $1579 \text{ cm}^{-1}$  and corresponds to C–H bending and C–N stretching

modes. The least intensive band in infrared spectra in water of inosine at  $1488 \text{ cm}^{-1}$  for anti- and  $1480 \text{ cm}^{-1}$  for syn-conformations of A is the most intensive band ( $1507 \text{ cm}^{-1}$ ) in the Raman spectrum. Our results from the theoretical IR spectrum of inosine in the gas phase show that the  $\delta_{\text{NH}}$  and  $\nu_{\text{C-N}}$  bands were found at  $1554 \text{ cm}^{-1}$  in the two conformers that corresponds with the experimentally observed line at  $1544 \text{ cm}^{-1}$  in the IR spectrum of inosine in KBr.

In the region below  $1500 \text{ cm}^{-1}$ , the inosine infrared spectrum in aqueous medium<sup>4</sup> shows three bands at  $1436$ ,  $1383$ , and  $1336 \text{ cm}^{-1}$ , which are characteristic ring vibrations. The latter two of them were also observed in the Raman spectra in water medium<sup>3</sup> and was calculated to be at  $1377$  and  $1346 \text{ cm}^{-1}$  in the case of the anti-conformer of A and  $1377$  and  $1352 \text{ cm}^{-1}$  for the syn-conformer. These bands were assigned to C–N and C–C stretching, C–C and C–H bending frequencies in IR spectra in water of the inosine two conformers of A.

In the experimental Raman spectrum of aqueous inosine, the ring breathing motion of inosine gives rise to an intense Raman line at  $721 \text{ cm}^{-1}$ , corresponding to that of the keto analogue but different from that of the enol analog at  $736 \text{ cm}^{-1}$ . According to our theoretical spectra this line was found at  $741 \text{ cm}^{-1}$  for anti-A and  $756 \text{ cm}^{-1}$  for anti-B. The ring breathing motion in the syn-conformer of A and B was predicted to be at  $697$  and  $682 \text{ cm}^{-1}$ .

Summing up, no band  $\nu_{\text{O-H}}$  was found in the experimental IR and Raman spectra of inosine in water for enol form B. According to theoretical spectra of the syn- and anti-conformers of tautomer B in water solution, this line was predicted to be shifted (because of hydrogen bonding formation with water molecules) around  $2920 \text{ cm}^{-1}$ . The C–H stretching frequency of the ribose ring was calculated to be around  $2930 \text{ cm}^{-1}$  and it overlaps with  $\nu_{\text{O-H}}$  one but the intensity of these lines is different. Therefore, the experimental IR and Raman spectra of inosine in water have no evidence of the enol tautomer presence but according to our theoretical



predictions with regard to kinetics of tautomeric conversion in water of inosine, the enol form **B** should be presented in solution. We have shown<sup>22,37–39</sup> that the calculated barriers in the range of 12–17 kcal mol<sup>−1</sup> for the water-assisted proton transfer reactions indicate that the tautomeric conversions are kinetically feasible processes. The detection of the rare tautomeric forms by spectroscopic methods is not possible and a probable exception might be fluorescence spectroscopy.<sup>37</sup>

#### 4. CONCLUSIONS

The presented study considers water-assisted proton transfer in the nucleoside inosine. MP2/6-31+G(d,p), SCS-MP2/6-31+G(d,p), CPCM/MP2/6-31+G(d,p), and CPCM/SCS-MP2/6-31+G(d,p) calculations show that in the gas phase and in water solution, the most stable is the amino-oxo tautomer **A** (6-keto form) followed by the imino-hydroxy tautomer **B1** (6-hydroxy form). The syn-conformer of tautomer **A** is more stable than the anti-conformer. The changing of the tautomeric form of the hypoxanthine moiety does not change the conformation of the nucleoside. The rate constant of the water-assisted proton transfer **A** → **B1** in inosine,  $k = 5.64 \times 10^1 \text{ s}^{-1}$ , is sufficiently large to generate imino-hydroxy tautomer. The time necessary to reach the equilibrium concentration of the 6-keto and 6-hydroxy tautomeric forms was calculated to be  $2.10 \times 10^{-6} \text{ s}$ .

Following the reaction path shown in Figure 3 it can be concluded that the solvent-assisted proton transfer in inosine is realized through the asynchronous concerted mechanism.

#### ■ ASSOCIATED CONTENT

##### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b11316.

MP2/6-31+G\*\* and SCS-MP2/6-31+G\*\* calculated energies of tautomers and conformers of inosine; CPCM/MP2/6-31+G\*\* and CPCM/SCS-MP2/6-31+G\*\* calculated energies of tautomers and conformers of inosine; CPCM/MP2/6-31+G\*\* and CPCM/SCS-MP2/6-31+G\*\* calculated energies of tautomers and conformers of inosine + 5H<sub>2</sub>O (PDF)

#### ■ AUTHOR INFORMATION

##### Corresponding Author

\*E-mail: venelin@orgchm.bas.bg. Phone: +3592-9606197.

##### ORCID

Venelin Enchev: 0000-0002-5481-3029

##### Notes

The authors declare no competing financial interest.

#### ■ ACKNOWLEDGMENTS

Funding of this work by the National Science Fund, under Grant DN09/7/2016 is gratefully acknowledged. The calculations were partially performed on the AVITOHOL supercomputer at the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences.

#### ■ REFERENCES

(1) Miles, H. T. Tautomeric forms in a polynucleotide helix and their bearing on the structure of DNA. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 791–802.

(2) Wolfenden, R. V. Tautomeric equilibria in inosine and adenosine. *J. Mol. Biol.* **1969**, *40*, 307–310.

(3) Medeiros, G. C.; Thomas, G. J., Jr. On the tautomeric structure of inosine. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1971**, *238*, 1–4.

(4) Psoda, A.; Shugar, D. Spectral studies on tautomeric forms of inosine. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1971**, *247*, 507–513.

(5) Evans, F. E.; Sarma, R. H. The tautomeric form of inosine in aqueous solution. *J. Mol. Biol.* **1974**, *89*, 249–253.

(6) Topal, M. D.; Fresco, J. R. Base pairing and fidelity in codon-anticodon interaction. *Nature* **1976**, *263*, 289–293.

(7) Subramanian, E. Inosine (hypoxanthine riboside) C10H12N4O5. *Cryst. Struct. Commun.* **1979**, *8*, 777–785.

(8) Munns, A. R. I.; Tollin, P. Crystal and molecular structure of inosine. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1970**, *26*, 1101–1113.

(9) Thewalt, U.; Bugg, C. E.; Marsh, R. E. The crystal structure of guanosine dihydrate and inosine dihydrate. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1970**, *26*, 1089–1101.

(10) Suzuki, Y.; Nagashima, N. Polymorphism of inosine. *Bull. Chem. Soc. Jpn.* **1970**, *43*, 1600.

(11) Subramanian, E.; Madden, J. J.; Bugg, C. E. A syn conformation for inosine, the wobble nucleoside in some tRNA's. *Biochem. Biophys. Res. Commun.* **1973**, *50*, 691–696.

(12) Chiarella, R. A.; Gillon, A. L.; Burton, R. C.; Davey, R. J.; Sadiq, G.; Auffret, A.; Cioffi, M.; Hunter, C. A. The nucleation of inosine: the impact of solution chemistry on the appearance of polymorphic and hydrated crystal forms. *Faraday Discuss.* **2007**, *136*, 179–193.

(13) Beetz, C. P., Jr.; Ascarielli, G. The low frequency vibrations of the nucleosides: Uridine, cytidine and inosine. Determination of vibrations associated with the ribose ring. *Spectrochim. Acta, Part A* **1980**, *36*, 525–534.

(14) Shishkin, O. V.; Pel'menschikov, A.; Hovorun, D. M.; Leszczynski, J. Molecular structure of free canonical 2'-deoxyribonucleosides: A density functional study. *J. Mol. Struct.* **2000**, *526*, 329–341.

(15) Brovarets, O. O.; Hovorun, D. M. Prototropic tautomerism and basic molecular principles of hypoxanthine mutagenicity: An exhaustive quantum-chemical analysis. *J. Biomol. Struct. Dyn.* **2013**, *31*, 913–926.

(16) Brovarets, O. O.; Yurenko, Y. P.; Dubey, I. Y.; Hovorun, D. M. Can DNA-binding proteins of replisome tautomerize nucleotide bases? Ab initio model study. *J. Biomol. Struct. Dyn.* **2012**, *29*, 1101–1109.

(17) Brovarets, O. O.; Hovorun, D. M. Stability of mutagenic tautomers of uracil and its halogen derivatives: The results of quantum-mechanical investigation. *Biopolym. Cell* **2010**, *26*, 295–298.

(18) Yurenko, Y. P.; Zhuravitsky, R. O.; Samijlenko, S. P.; Ghomi, M.; Hovorun, D. M. The whole of intramolecular H-bonding in the isolated DNA nucleoside thymidine. AIM electron density topological study. *Chem. Phys. Lett.* **2007**, *447*, 140–146.

(19) Mejía-Mazariegos, L.; Robles, J.; García-Revilla, M. A. Tautomerism in some pyrimidine nucleoside analogues used in the treatment of cancer: an ab initio study. *Theor. Chem. Acc.* **2016**, *135*, 233.

(20) Alvarez-Ros, M. C.; Palafox, M. A. Molecular structure of the nucleoside analogue inosine using DFT methods: Conformational analysis, crystal simulations and possible behaviour. *J. Mol. Struct.* **2013**, *1047*, 358–371.

(21) Nir, E.; Plützer, C.; Kleinermanns, K.; de Vries, M. Properties of isolated DNA bases, base pairs and nucleosides examined by laser spectroscopy. *Eur. Phys. J. D* **2002**, *20*, 317–329.

(22) Markova, N.; Pejov, L.; Stoyanova, N.; Enchev, V. Hybrid MC/QC simulations of water-assisted proton transfer in nucleosides. Guanosine and its analog acyclovir. *J. Biomol. Struct. Dyn.* **2017**, *35*, 1168–1188.

(23) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular



orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

(24) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(25) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681.

(26) Gonzalez, C.; Schlegel, H. B. An improved algorithm for reaction path following. *J. Chem. Phys.* **1989**, *90*, 2154–2161.

(27) Grimme, S. Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(28) Galvao, T. L. P.; Rocha, I. M.; da Silva, M.; da Silva, M. From 2-hydroxypyridine to 4(3H)-pyrimidinone: Computational study on the control of the tautomeric equilibrium. *J. Phys. Chem. A* **2013**, *117*, 12668–12674.

(29) Gordon, M. S.; Schmidt, M. W. *Advances in Electronic Structure Theory: GAMESS a Decade Later*; Elsevier, 2005; pp 1167–1189.

(30) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; et al. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(31) Saenger, W. *Principles in Nucleic Acid Structure*; Springer Verlag: New York, 1984.

(32) Lord, R. C.; Thomas, G. J., Jr. Raman spectral studies of nucleic acids and related molecules-I Ribonucleic acid derivatives. *Spectrochim. Acta, Part A* **1967**, *23*, 2551–2591.

(33) Peng, C. S.; Tokmakoff, A. Identification of lactam-lactim tautomers of aromatic heterocycles in aqueous solution using 2D IR spectroscopy. *J. Phys. Chem. Lett.* **2012**, *3*, 3302–3306.

(34) Singh, V.; Fedeles, B. I.; Essigmann, J. M. Role of tautomerism in RNA biochemistry. *RNA* **2015**, *21*, 1–13.

(35) Bell, A. F.; Hecht, L.; Barron, L. D. Vibrational Raman optical activity of pyrimidine nucleosides. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 553–562.

(36) Heidari, A. An Analytical and Computational Infrared Spectroscopic Review of Vibrational Modes in Nucleic Acids. *Austin J. Anal. Pharm. Chem.* **2016**, *3*, No. 1058.

(37) Markova, N.; Enchev, V.; Timtcheva, I. Oxo-hydroxy tautomerism of 5-fluorouracil: water-assisted proton transfer. *J. Phys. Chem. A* **2005**, *109*, 1981–1988.

(38) Markova, N.; Enchev, V.; Ivanova, G. Tautomeric equilibria of 5-fluorouracil anionic species in water. *J. Phys. Chem. A* **2010**, *114*, 13154–13162.

(39) Markova, N.; Pejov, L.; Enchev, V. A hybrid statistical mechanics - Quantum chemical model for proton transfer in 5-azauracil and 6-azauracil in water solution. *Int. J. Quantum Chem.* **2015**, *115*, 477–485.

(40) Spectral Database for Organic Compounds; National Institute of Advanced Industrial Science and Technology (AIST): Japan. <http://sdb.sdb.aist.go.jp> (accessed Aug 7, 2018).