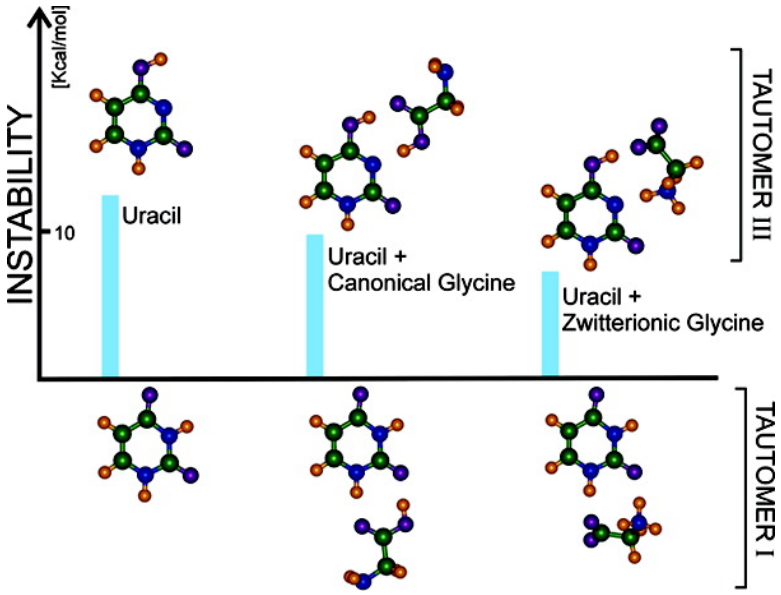
Article

Interaction with Glycine Increases Stability of a Mutagenic Tautomer of Uracil. A Density Functional Theory Study

Iwona Da#bkowska, Maciej Gutowski, and Janusz Rak

J. Am. Chem. Soc., **2005**, 127 (7), 2238-2248• DOI: 10.1021/ja048730k • Publication Date (Web): 29 January 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 01/29/2005

Interaction with Glycine Increases Stability of a Mutagenic Tautomer of Uracil. A Density Functional Theory Study

Iwona Dąbkowska,[†] Maciej Gutowski,*[,‡] and Janusz Rak*[,†]

Contribution from the Department of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland, and Chemical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352

Received March 5, 2004; E-mail: maciej.gutowski@pnl.gov; janusz@raptor.chem.univ.gda.pl

Abstract: The most stable structures for the gas-phase complexes of minor tautomers of uracil (U) with glycine (G) were characterized at the density functional B3LYP/6-31++G** level of theory. These are cyclic structures stabilized by two hydrogen bonds. The relative stability of isolated tautomers of uracil was rationalized by using thermodynamic and structural arguments. The stabilization energies for complexes between the tautomers of U and G result from interplay between the stabilizing two-body interaction energies and destabilizing one-body terms. The latter are related to the energies of (i) tautomerization of the unperturbed moieties and (ii) distortions of the resulting rare tautomers in the complex. The two-body term describes the interaction energy between distorted tautomers. The two-body interaction energy term correlates with perturbations of length of the proton-donor bonds as well as with deprotonation enthalpies and proton affinities of the appropriate monomer sites. It was demonstrated that the relative instability of rare tautomers of uracil is diminished due to their interactions with glycine. In particular, the instability of the third most stable tautomer (U^{III}) is decreased from 11.9 kcal/mol for non-interacting uracil to 6.7 kcal/mol for uracil in a complex with the zwitterionic tautomer of glycine. A decrease of instability by 5.2 kcal/mol could result in an increase of concentration of U^{III} by almost 5 orders of magnitude. This is the tautomer with proton donor and acceptor sites matching guanine rather than adenine. Moreover, kinetic characteristics obtained for the glycine-assisted conversion of the most stable tautomer of uracil (U^{I}) to U^{III} indicate that the $U^{I}\leftrightarrow U^{III}$ thermodynamic equilibrium could be easily attained at room temperature. The resulting concentration of this tautomer falls in a mutationally significant range.

I. Introduction

Point mutations, which develop during replication of DNA or RNA, may result from the occurrence of rare imino/enol tautomeric forms of nucleic bases (NBs) in physiological conditions.[1] According to Löwdin,[2] point mutations can occur via concerted transfer of two protons in the Watson−Crick-type dimer of NBs, with the product being a dimer formed by two rare tautomers of NBs. For the adenine-thymine (AT) pair such a process would lead to a thymine tautomer, T^{III}, which can preferentially bind to guanine rather than to adenine, leading directly to the formation of a mispair of NBs (see Figure 1). The formation of T^{III} is rather improbable within the complementary AT base pair (i.e., in DNA), as was shown by the groups of Hobza[3,4] and Leszczyński.[5] The tautomerization can, however, occur in physiological conditions where interactions with other cellular components (e.g., amino acids or ions) could stabilize the T^{III} form.

The calculated dipole moments (μ) of tautomers of a NB can differ significantly; e.g., the B3LYP/6-31+G(d,p) values of μ span a range from 1.3 to 9.4 D for the tautomers of uracil.[6] Hence, the relative stability of tautomers can be affected by interactions with polar environments. Moreover, there could be a wide range of specific intermolecular interactions leading to an extra stabilization of various forms of NBs. For example, N4-imino-1-methylcytosine was found to be stabilized by coordination of N4 to a platinum(IV) ion,[7] and cis-platinum binding to the N7 site of 9-ethylguanine was shown to facilitate ionization of its N1 proton.[8]

In this contribution, we explore the effect of intermolecular interactions on the tautomeric equilibrium of uracil. We select glycine as a model system which may form cyclic hydrogen bonds with uracil and which may exist in either canonical (G^C) or zwitterionic structure (G^Z) (see Figure 2), and depending on the environment, either the former or the latter form becomes dominant. The canonical and zwitterionic forms of glycine differ in polarity. Thus they provide different electrostatic environments for various tautomers of uracil.

[†] University of Gdańsk.
[‡] Pacific Northwest National Laboratory.

(1) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285−289.
(2) Löwdin, P. O. *Adv. Quantum Chem.* **1965**, *2*, 213−360.
(3) Hrouda, V.; Florian, J.; Hobza, P. *J. Phys. Chem.* **1993**, *97*, 1542−1557.
(4) Florian, J.; Hrouda, V.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 1457−1460.
(5) Florian, J.; Leszczyński, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010−3017.
(6) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 1288−1295 and references therein.
(7) Lippert, B.; Schollhorn, H.; Thewalt, U. *J. Am. Chem. Soc.* **1986**, *108*, 6616−6621.
(8) Lippert, B. *J. Am. Chem. Soc.* **1981**, *103*, 5691−5697.

10.1021/ja048730k CCC: $30.25

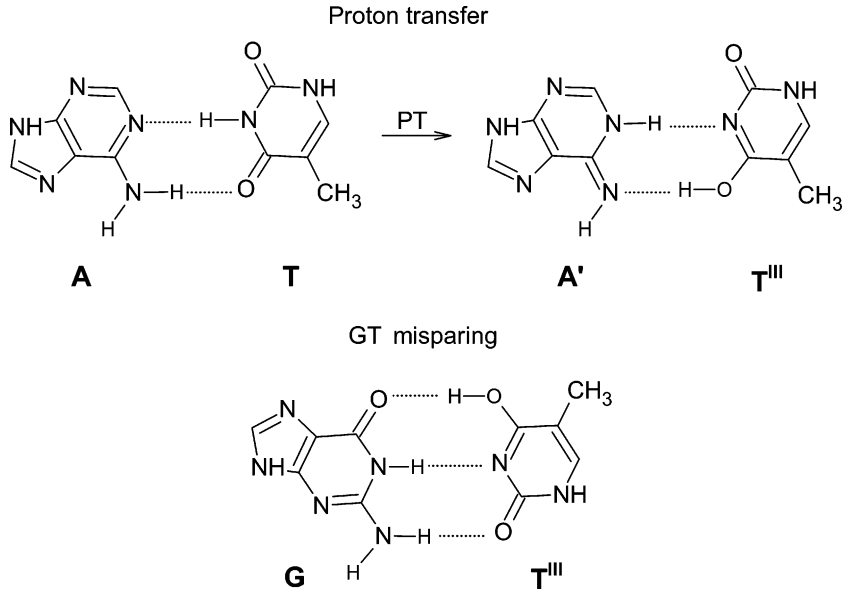


Figure 1. A rare tautomer of uracil, U^{III}, can mimic cytosine in the hydrogen-bonded complex with guanine.

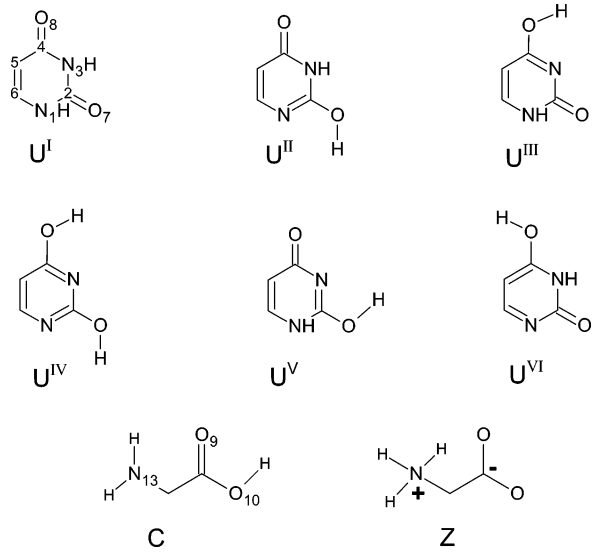


Figure 2. Tautomers of uracil and glycine.

In the gas phase, only G^C is present, as concluded from its microwave spectrum.[9,10] The same conclusion was reached from the analysis of vibrational spectra of glycine in rare gas matrices and helium clusters.[11] These experimental findings are consistent with computational results,[12−16] which indicate that G^Z is not even a minimum on the potential energy surface. However, intermolecular interactions can affect the canonical/zwitterion equilibrium. For example, it was computationally predicted that two water molecules are sufficient to render the zwitterion geometrically stable,[16] though energetically unstable. Furthermore, for glycine solvated by three water molecules, configurations based on the zwitterion are almost equienergetic with those involving the canonical form.[17] Recent results from Bowen's group revealed that five water molecules are necessary to observe a zwitterion-like photoelectron spectroscopy peak of anionic glycine.[18] Finally, in water solutions, over a wide range of pH, glycine exists primarily in its zwitterionic form.[19,20]

Theoretical and experimental studies also explored agents other than water which can enhance the stability of zwitterions in the gas phase.[21,22] Of particular interest were interactions between amino acids and protons or alkali metal cations, as the extra positive charge could stabilize the zwitterion structure through an intramolecular salt bridge.[21,22] Indeed, Bowers et al. suggested that singly and doubly N-methylated glycine forms a salt bridge structure independent of the metal cation (Na^+ or Rb^+), whereas doubly C-methylated glycine forms a charge solvation structure (canonical form) when rubidiated and a salt bridge when sodiated.[22]

An excess electron was considered as another species which can enhance the stability of zwitterions in the gas phase.[23−25] The neutral zwitterionic tautomer is expected to have a larger dipole moment than the canonical tautomer. Therefore, formation of a dipole-bound anionic state was expected to stabilize the zwitterion with respect to the canonical form. Indeed, recent theoretical results indicate that G^Z becomes a local minimum upon attachment of an excess electron, but the dipole-bound anion based on the canonical structure remains still more stable than this based on the zwitterionic structure.[23] The computational results for anionic zwitterions of glycine,[23] betaine,[25] and

(9) Suernam, R. D.; Lovas, F. J. *J. Mol. Spectrosc.* **1980**, *72*, 372−382.
(10) Lovas, F. J.; Kawashima, Y.; Grabov, J. U.; Suenram, R. D.; Fraser, G. T.; Hirota, E. *Astrophys. J.* **1995**, *455*, L201−L204.
(11) Huisken, F.; Werhahn, O.; Ivanov, A. Y.; Krasnokutsi, S. A. *J. Chem. Phys.* **1999**, *111*, 2978−2984.
(12) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917−7924.
(13) Ding, Y.; Krogh-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261−266.
(14) Yu, D.; Armstrong, D. A.; Rauk, A. *Can. J. Chem.* **1992**, *70*, 1762−1772.
(15) Csaszar, A. G. *J. Am. Chem. Soc.* **1992**, *114*, 9568−9575.
(16) Gordon, M. S.; Jensen, J. H. *Acc. Chem. Res.* **1996**, *29*, 536−543.
(17) Kassab, E.; Langlet, J.; Evleth, E.; Akacem, Y. *J. Mol. Struct. (THEOCHEM)* **2000**, *531*, 267−282.
(18) Xu, S.; Nilles, J. M.; Bowen, K. H. *J. Chem. Phys.* **2003**, *119*, 10696−10701.
(19) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry*; W.H. Freeman and Co.: New York, 1980; Part I, pp 41−53.
(20) Tortonda, F. R.; Pascual-Ahuir, J. L.; Silla, E.; Tuñón, I. *Chem. Phys. Lett.* **1996**, *260*, 21−26.
(21) Wyttenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *183*, 243−252.
(22) Wyttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458−3464.
(23) Gutowski, M.; Skurski, P.; Simons, J. *J. Am. Chem. Soc.* **2000**, *122*, 10159−10162.
(24) Skurski, P.; Rak, J.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11073−11074.
(25) Rak, J.; Skurski, P.; Gutowski, M. *J. Chem. Phys.* **2001**, *114*, 10673−10681.

arginine[24] indicate that the excess electron is bound vertically by 6.5−9.1 kcal/mol.

Uracil (U) and thymine (T) are pyrimidine bases occurring in RNA and DNA, respectively. U also serves as a model of T, its 5-methyl derivative illustrated in Figure 1. U in RNA and T in DNA are bonded to the sugar−phosphate backbone through the nitrogen N1 (see Figure 2). Thus, the N1 site cannot participate in tautomerizations in nucleotides of uracil and thymine. Uracil in the gas phase may exist in six tautomeric forms due to the presence of four proton acceptor and donor sites (O7, O8, N1, N3; see Figure 2). In addition, every hydroxy tautomer of U possesses rotamers resulting from two possible orientations of each OH group.[6,26]

The tautomers of uracil have been the subject of numerous experimental and theoretical studies.[6,26] NMR, UV, IR, Raman, and microwave spectroscopic experiments indicate that 2,4-dioxouracil (U^{I} in Figure 2) is the most stable tautomer in the gas and solid phases as well as in solutions.[6,26] Calorimetric measurements, carried out by Beak et al.,[27] demonstrated that the 2-oxo-4-hydroxy tautomer of uracil (U^{III} in Figure 2) is less stable than the 2,4-dioxo form by 19 ± 6 kcal/mol. Furthermore, Fuji and co-workers[28,29] concluded, by using UV technique, that the most stable minor tautomer of uracil is less stable than the most abundant form by ca. 9.6 kcal/mol.

Theoretical predictions are in good agreement with the experimental findings. The six most stable tautomers of uracil were studied at an effective QCISD/6-311+G(2d,p) level by Wolken and Tureček.[30] Their results suggest that in the gas-phase equilibrium mixture at 523 K, the content of the most stable dioxo tautomer is >99.9%. The diketo tautomer is stable by ca. 11 and 12 kcal/mol with respect to the lowest monohydroxy and dihydroxy tautomer, respectively. Moreover, the most stable rare tautomer, 2-hydroxy-4-oxo (U^{II} in Figure 2), is more stable than the 2-oxo-4-hydroxy form (U^{III} in Figure 2) by <1 kcal/mol. A satisfactory agreement was found between the B3LYP and QCISD relative energies. The most thorough computational study of uracil has recently been completed by Kryachko et al. at the B3LYP/6-31+G(d,p) level.[6] Their findings were consistent with those of Wolken and Tureček and covered in addition some rotamers of uracil.

In RNA/DNA the N1 position of uracil/thymine cannot be involved in tautomerizations because the base is covalently attached to a sugar through this atom. Consequently, the U^{III} tautomer, which is unstable with respect to U^{I} by 12 kcal/mol,[30] gains significance over the U^{II} and U^{IV} tautomers. If we assume that the same instability holds for uracil in the uridine triphosphate, then at room temperature the concentration of the U^{III} tautomeric form should be 10^9 times smaller than that of the canonical nucleotide. In addition, if a damaged nucleotide was incorporated into a growing DNA/RNA strand, it could be easily excised by the DNA/RNA polymerases through their proofreading activity,[31] providing that the nucleotide mismatches the template base or exhibits an unusual structure. It seems, therefore, that the mechanism of point mutations based on the equilibrium between nucleotides in different tautomeric forms[1] cannot be practically relevant without additional stabilization of rare tautomers in their cellular environment.

The possibility of stabilization of rare tautomers of uracil due to complexation with one water molecule has recently been studied by Kryachko et al.[32] They found that the relative instability of rare tautomers with respect to the diketo form decreases by no more than 3 kcal/mol as a result of the interaction with one water molecule, though the relative ordering of tautomers remains unchanged. The interactions of tautomers of the DNA/RNA bases with Na^+ and K^+ have been investigated by Russo et al. at the B3LYP/6-311+G(2df,2p) level.[33] Those authors demonstrated that the relative instability of the U^{III} tautomer (see Figure 2) bound in a complex with an alkali metal cation diminishes, in comparison with the instability of an isolated U^{III}, by as much as 8.5 and 6.6 kcal/mol for Na^+ and K^+, respectively.

The main goal of the present investigation is to explore whether hydrogen bonding between nucleic bases and polar molecules could change the relative stability of tautomers of NBs. For this purpose we shall use the uracil-glycine dimer, and we will study complexes of different tautomers of uracil with glycine in its canonical and zwitterionic forms. We will demonstrate that the largest stabilization occurs for the U^{III} tautomer bound to the zwitterionic glycine. The increased stability of the 2-oxo-4-hydroxy tautomer of uracil, and probably thymine, can be relevant to point mutations because its proton donor and acceptor sites match guanine rather than adenine (see Figure 1).

The stabilization energies for complexes between the tautomers of U and G will be decomposed into two-body interaction energies and one-body terms. The latter are related to the energy of (a) tautomerization of the unperturbed moieties and (b) distortion of the resulting rare tautomers in the complex. We will then explore correlations between the two-body interaction energies and (a) modifications of proton donor bond lengths and (b) proton affinities and deprotonation enthalpies of the centers involved in hydrogen bonds.

II. Computational Method

This computational effort is a continuation of our previous studies on the neutral complexes between the most stable tautomers of uracil and glycine[34,35] as well as on the anionic species.[36] Here we will use the notation U^RG^Ln, where $R = $ I, ..., VI denotes the Rth most stable tautomer of uracil, L becomes Z (C) for the zwitterionic (canonical) form of glycine, and n designates the nth most stable complex between U^R and G^L. Thus, $U^{III}G^Z2$ indicates the second most stable complex formed between the third most stable tatutomer of uracil and the zwitterionic glycine.

The minimum energy structures for complexes involving the U^R and G^L moieties were initially identified using a computationally efficient semiempirical method, PM3.[37] To obtain the final results, we applied

(26) Tian, S. X.; Zhang, Z. J.; Chen, X. J.; Xu, K. Z. *Chem. Phys.* **1999**, *242*, 217 and references therein.
(27) Beak, P.; White, J. M. *J. Am. Chem. Soc.* **1982**, *104*, 7073−7077.
(28) Fujii, M.; Tamura, T.; Mikami, N.; Ito, M. *Chem. Phys. Lett.* **1986**, *126*, 583−587.
(29) Tsuchiya, Y.; Tamura, T.; Fujii, M.; Ito, M. *J. Phys. Chem.* **1988**, *92*, 1760−1765.
(30) Wolken, J. K.; Tureček, F. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1065−1071.
(31) Beckman, R. A.; Loeb, L. A. *Q. Rev. Biophys.* **1993**, *26*, 225−331.
(32) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 1934−1943.
(33) Russo, N.; Toscano, M.; Grand, A. *J. Am. Chem. Soc.* **2001**, *123*, 10272−10279.
(34) Dabkowska, I.; Gutowski, M.; Rak, J. *Pol. J. Chem. A* **2002**, *76*, 1243−1247.
(35) Dabkowska, I.; Rak, J.; Gutowski, M. *J. Phys. Chem. A* **2002**, *106*, 7423−7433.
(36) Gutowski, M.; Dabkowska, I.; Rak, J.; Xu, S.; Nilles, J. M.; Radisic, D.; Bowen, K. H., Jr. *Eur. Phys. J. D* **2002**, *20*, 431−439.
(37) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209; 221−264.

primarily the DFT method with a hybrid B3LYP functional[38−40] and 6-31++G** (6d) basis sets.[41] Our recent work on the neutral and charged arginine,[24,42] as well as other reports on complexes between nucleic acid bases and water,[32,43] and pyridine and water,[44] demonstrated the usefulness of this approach in studying systems with intra- and intermolecular hydrogen bonds.

For some U^RG^Z complexes we observed a barrier-free double proton transfer at the B3LYP level of theory. The B3LYP method, however, is known to underestimate activation barriers in comparison with experimental values or with the results of accurate ab initio calculations.[45−47] For this reason, some of these barrier-free double proton transfers could be artifacts of our computational approach. The second-order Møller−Plesset method (MP2) with the same 6-31++G** (6d) basis set was employed to determine whether the double-proton-transfer processes are indeed barrier-free. The core 1s orbitals of C, N, and O were excluded from electron correlation treatments at the MP2 level.

The stability of a complex U^RG^Ln is expressed in terms of E_{stab}, H_{stab}, and G_{stab}. E_{stab} is defined as the energy of U^RG^Ln measured with respect to the energies of the most stable tautomers of U and G, i.e., U^I and G^C,

$$E_{stab} = E^{U^I||U^I} + E^{G^C||G^C} - E^{U^RG^Ln||U^RG^Ln} \quad (1)$$

where $E^{A||B}$ denotes the energy of A at the optimal geometry of B. Thus a positive value of E_{stab} is a measure of the strength of the hydrogen bonds. The stabilization enthalpy H_{stab} results from correcting E_{stab} for zero-point vibration terms, thermal contributions to energy from vibrations, rotations, and translations, and the pV terms. Finally, the stabilization Gibbs energy G_{stab} results from supplementing H_{stab} with the entropy term. These terms were calculated in the rigid rotor−harmonic oscillator (RRHO) approximation. The values of H_{stab} and G_{stab} discussed in section III were obtained for $T = 298$ K and $p = 1$ atm.

The stabilization energy E_{stab} for U^RG^Ln is decomposed as

$$E_{stab} = E^U_{1-b} + E^G_{1-b} + E_{2-b} \quad (2)$$

where $E_{1-b}{}^X$ is a repulsive one-body component related to a tautomerization and distortion of the monomer X (X = U, G) in the dimer,

$$E^U_{1-b} = E^{U^I||U^I} - E^{U^R||U^RG^Ln} \quad (3)$$

$$E^G_{1-b} = E^{G^C||G^C} - E^{G^L||U^RG^Ln} \quad (4)$$

and E_{2-b} is a two-body interaction energy between the distorted tautomers U^R and G^L in the U^RG^Ln complex,

$$E_{2-b} = E^{U^R||U^RG^Ln} + E^{G^L||U^RG^Ln} - E^{U^RG^Ln||U^RG^Ln} \quad (5)$$

The one-body term $E_{1-b}{}^U$ is further split into the energy of tautomerization for an isolated uracil and the deformation energy of the U^R moiety in the U^RG^Ln complex:

$$E^U_{1-b} = E^{U-tau}_{1-b} + E^{U-def}_{1-b} \quad (6)$$

$$E^{U-tau}_{1-b} = E^{U^I||U^I} - E^{U^R||U^R} \quad (7)$$

$$E^{U-def}_{1-b} = E^{U^R||U^R} - E^{U^R||U^RG^Ln} \quad (8)$$

The zwitterion of glycine, on the other hand, is not a minimum in the gas phase.[23] Hence, the E^G_{1-b} term is defined with respect to the energy of canonical glycine and is not further split into the tautomerization and deformation components.

The acidity and basicity of isolated monomers were characterized by the deprotonation enthalpy (DPE) and proton affinity (PA), respectively, which were determined in the RRHO approximation. More specifically, the symbol DPE(*n*;*m*) (PA(*n*;*m*)) denotes the deprotonation enthalpy (proton affinity) of the *n*th tautomer, and *m* indicates which atom is involved in the proton detachment (attachment) process.

The basis set and methodological saturation tests were performed for complexes of the most stable tautomers of U and G.[35] The values of E_{stab} obtained with the 6-31++G**[41] and aug-cc-pVDZ[48] basis sets differed by less than 0.4 kcal/mol, and so did the one- and two-body terms. The discrepancies between the MP2 and B3LYP values of E_{stab} did not exceed 1.2 kcal/mol, and geometrical parameters of hydrogen bonds differed by less than 0.03 Å for bond lengths and 1° for bond angles. Moreover, the relative instabilities of the U^{III} and U^V tautomers, calculated recently by Russo et al.[33] at the B3LYP/6-311+G(2df,2p) level, differ from our results by less than 0.6 kcal/mol. Hence, in the current study we employed primarily the B3LYP/6-31++G** approach. Moreover, the value of basis set superposition error[49] (BSSE) did not exceed 1.5 kcal/mol in the U^IG^Cn family of complexes.[35] In the current study, the values of E_{2-b} are typically 1 order of magnitude larger than typical values of BSSE. Hence, our results are presented without the counterpoise corrections.[49]

Hydrogen bonding can be studied in detail for small systems using highly correlated electronic structure methods and symmetry-adapted perturbation theory of intermolecular forces.[50] One also needs qualitative interpretations that can be used for large systems or for a variety of systems.[51] Simple, robust, and general potentials, which are based on monomer properties and enable predictions of the stability of hydrogen-bonded complexes, would greatly simplify time-consuming ab initio explorations of potential energy surfaces and could be applied in designing and modeling of supramolecular systems.

The most common qualitative approach is to relate the strength of a hydrogen bond to the values of PA and DPE for the proton acceptor and donor, respectively. Every uracil-glycine complex studied here possesses two hydrogen bonds, and we will use the variables $DPE^G - PA^U \equiv x1$, and $DPE^U - PA^G \equiv x2$ to characterize the propensity of proton donor and acceptor pairs of uracil and glycine to form hydrogen bonds. In addition, we will perform least-squares fitting of the E_{2-b} term to various polynomials of *x1* and *x2*.

Tautomerization reactions have to fulfill not only thermodynamic but also certain kinetic limits to be relevant to spontaneous DNA mutations.[4,5] First, the lifetime of the canonical base should be shorter than the reproduction period of a given species. Second, the rare tautomer needs to remain stable during the time period from the occurrence of tautomerization until the replication process is completed. These conditions impose constraints on barriers for the forward and reverse reactions of tautomerization. We will concentrate on the reaction barriers for the $U^I \leftrightarrow U^{III}$ tautomerization because the forward process

(38) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(40) Lee, C.; Yang, W.; Paar, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(41) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(42) Rak, J.; Skurski, P.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11695−11707.
(43) Chandra, A. K.; Nguyen, M. T.; Uchimaru, T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1999**, *103*, 8853−8860.
(44) Dkhissi, A.; Adamowicz, L.; Maes, G. *J. Phys. Chem. A* **2000**, *104*, 2112−2119.
(45) Jursic, B. S. In *Theoretical and Computational Chemistry*; Seminario, J. M., Ed.; Elsevier: Amsterdam, 1996; Vol. 4, p 709.
(46) Sadhukhan, S.; Munoz, D.; Adamo, C.; Scuseria, G. E. *Chem. Phys. Lett.* **1999**, *306*, 83−87.
(47) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936−2941.
(48) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(49) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.
(50) Chalasiński, G.; Szczesniak, M. M. *Chem. Rev.* **2000**, *100*, 4227−4252.
(51) Abraham, M. H.; Platts, J. A. *J. Org. Chem.* **2001**, *66*, 3484−3491.

Table 1. Tautomerization Enthalpies, ΔH_{tau} (kcal/mol), for the $U^I \rightarrow U^R$ Reactions Decomposed into Formal Steps Related to Proton Detachment from Appropriate Sites of U^I and U^R (for Structures of Tautomers, See Figure 2)

tautomer-ization	formal decomposition[a]	numerical values for formal steps	ΔH_{tau}
$U^I \rightarrow U^{II}$	DPE(U^I;N1) − DPE(U^{II};O7)	332.8−321.9	10.7
$U^I \rightarrow U^{III}$	DPE(U^I;N3) − DPE(U^{III};O8)	346.2−334.5	11.7
$U^I \rightarrow U^{IV}$	DPE(U^I;N1) − DPE(U^{II};O7) + DPE(U^{II};N3) − DPE(U^{IV};O8)	332.8−321.9 + 337.3−335.6	12.6
$U^I \rightarrow U^V$	DPE(U^I;N3) − DPE(U^V;O7)	346.2−327.0	19.2
$U^I \rightarrow U^{VI}$	DPE(U^I;N1) − DPE(U^{VI};O8)	332.8−312.0	20.8

[a] The symbol DPE(n;m) denotes the deprotonation enthalpy of the nth tautomer, and m indicates which atom is involved in the proton detachment process. The values of DPE are calculated at the B3LYP/6-31++G** level.

Table 2. Proton Affinities (PA) of the O and N Atoms and Deprotonation Enthalpies (DPE) of the NH and OH Bonds Obtained at the B3LYP/6-31++G** Level for Selected Sites of Tautomers of Uracil and Glycine (All Results in kcal/mol)

	PA		DPE	
tautomer	site	value	site	value
U^I	O7	194.7	N1−H	332.8
U^{II}	N1	205.6	O7−H	321.9
	O8 (C5 side)	210.3	C5−H	386.7
	O8 (N3 side)	206.0	N3−H	337.6
U^{III}	N3	214.4	O8−H	334.5
U^{IV}	N1	211.3	O7−H	338.7
	N3	207.8	O8−H	335.6
U^V	N3	215.3[a]	O7−H	327.0
U^{VI}	N1	226.4	O8−H	335.6
	O7	211.0	N3−H	337.1
G^C	O9	183.4	O10−H	340.2
G^Z	O9	235.7	N13−H	317.0

[a] There is a discrepancy of 12 kcal/mol between the value reported here and that from ref 6. We believe the latter is erroneous.

can lead to mispairing with guanine (see Figure 1). The time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentration of U^{III} in the system of reversible first-order forward (k) and reverse (k') reactions ($U^I \leftrightarrow U^{III}$) can be estimated from[52]

$$\tau_{99.9\%} = \frac{\ln(10^3)}{k + k'} \tag{9}$$

and the half-life, $t_{1/2}$, of U^{III} is given by $\ln 2/k'$. We apply the standard transition state theory to estimate the values of k and k'.[52]

All calculations were carried out with the MOPAC 2000,[53] GAUSSIAN 98,[54] and NWChem,[55] codes on a cluster of 32-bit Xeon/SCI Dolphin processors, a cluster of 2xPentium III processors, IBM SP/2, and SGI Origin2000 numerical servers.

III. Results and Discussion

A. Properties of Monomers. Uracil could formally exist in six tautomeric forms, whereas glycine could exist in only two (see Figure 2). All tautomers of uracil, except for the 2,4-dioxo isomer U^I, may additionally form rotamers resulting from two orientations of each hydroxyl group. Hence, two rotamers correspond to the U^{II}, U^{III}, U^V, and U^{VI} isomers, while four correspond to the U^{IV} structure. Thus, one can distinguish 13 isomers that differ in the position of hydrogen atoms or/and the orientation of hydroxyl groups.[6]

The enthalpy of tautomerization, ΔH_{tau}, can be rigorously expressed as a linear combination of deprotonation enthalpies of appropriate neutral tautomers (or as a linear combination of proton affinities of appropriate neutral tautomers). For example, the enthalpy of the $U^I \rightarrow U^{III}$ reaction is given as either DPE(U^I;N3) − DPE(U^{III};O8) or PA(U^I;O8) − PA(U^{III};N3) (see Tables 1 and 2).

The tautomers of uracil are arranged in Figure 2 in the order of decreasing stability, and the values of ΔH_{tau}, calculated with respect to U^I, are presented in Table 1. Apparently, the dihyodroxy U^{IV} tautomer is more stable than the two monodydroxy tautomers, U^V and U^{VI}, due to aromatic stabilization (see Figure 2). By analyzing the DPE(n;m) values for monohydroxy tautomers, one can rationalize their relative stabilities. U^{II} is more stable than U^{VI} by 10.1 kcal/mol, which results from the second deprotonation step: DPE(U^{II};O7) and DPE(U^{VI};O8) for U^{II} and U^{VI}, respectively (see Table 1). The latter value is smaller than the former due to the repulsive interaction, absent in U^{II}, between the hydroxyl hydrogen of U^{VI} and the C5−H hydrogen. Similarly, U^{III} is more stable than U^V by 7.5 kcal/mol, and the difference results from the second deprotonation step, which is DPE (U^{III};O8) for U^{III} and DPE(U^V;O7) for U^V (see Table 1). The latter value is smaller than the former because the double bonds in U^V are isolated, whereas they are conjugated in U^{III} (see Figure 2).

The ΔH_{tau} values for U^{II} and U^{III} tautomers of 10.7 and 11.7 kcal/mol, respectively, are very similar (see Table 1), in good agreement with earlier theoretical predictions.[6] Both molecules possess a system of two conjugated double bonds in the ring, C═C and C═N, and in both species an attractive interaction exists between a hydrogen atom of the hydroxyl group and the lone electron pair of a heterocyclic nitrogen. U^{II}, however, is somewhat more stable than U^{III}. This may be related to a different number of intramolecular hydrogen bonds D−H···X (D, X = O or N) in U^{II} and U^{III}. There are three such interactions in U^{II} (N3−H···O(7), N3−H···O(8), and O7−H···N1) but only two, N1−H···O(7) and O8−H···N3, in U^{III} (see Figure 2).

In Figure 2, only the lowest energy rotamer is depicted for every tautomer. It is relatively easy to decide which of the two or four rotamers is the most stable. As far as monohydroxy structures are concerned, the most stable rotamer is that in which the hydroxyl group is capable of interacting with the lone electron pair of the neighboring heterocyclic nitrogen. In the case of U^{VI}, such an interaction cannot take place, and the OH group in the more stable rotamer is oriented toward a hydrogen atom, which is less acidic. The energy differences between rotamers in the monohydroxy tautomers span a range of 3−11 kcal/mol.[6] The O−H···N interactions are also responsible for the relative stability of dihydroxy rotamers. Moreover, it is preferable to engage each heterocyclic nitrogen in one O−H···N interaction rather than to have a single nitrogen atom involved in two O−H···N interactions. Thus, a rotamer of U^{IV} displayed in Figure 2 is the most stable. The energy differences between rotamers in the dihydroxy tautomer span a range of 1−5 kcal/mol.[6]

Glycine is much simpler than uracil since only two tautomers are involved, i.e., canonical (C) and zwitterionic (Z) (see Figure 2). The geometry optimization for G^Z in the gas phase ends up at G^C.[23] One may, however, estimate the gas-phase instability

(52) Attkins, P. W. *Physical Chemistry*; Oxford University Press: England, 1998.
(53) Stewart, J. J. P. MOPAC2000, Fujitsu Ltd., 1999.
(54) Frisch, M. J.; et al. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(55) Harrison, R. J., et al. NWChem, A Computational Chemistry Package for Parallel Computers, Version 4.0.1; Pacific Northwest National Laboratory: Richland, WA, 2001.

of G^Z by assuming an equilibrium structure of G^Z from solution. Our structure of G^Z has been optimized at the B3LYP/6-31++G** level in water solution, described by the PCM self-consistent reaction field model.[56] This structure leads to the gas-phase instability of G^Z of 23.7 kcal/mol at the B3LYP/6-31++G** level. A similar instability of 18.3 kcal/mol has been predicted at the coupled cluster level of theory, based on a structure of G^Z stabilized by an excess electron.[23] In addition, other minima exist on the potential energy surface of G^C due to a relatively "free" rotation around the C−N and C−C bonds, and two possible orientations of the OH group. For example, it was demonstrated computationally that the lowest energy canonical conformer is separated by less than 1 kcal/mol from another conformer in which OH acts as a proton donor and the lone electron pair of nitrogen acts as a proton acceptor.[23]

B. Geometries of U^RG^Ln's. Six tautomers of uracil and two tautomers of glycine can form 12 families of mixed dimers. We focused on the most stable complexes in every family, because they should be the most relevant in experimental conditions. Any tautomeric form of glycine and uracil possesses both proton acceptor and donor sites. Hence, we studied cyclic complexes, in which uracil and glycine are connected via two hydrogen bonds. Formation of complexes with three strong hydrogen bonds is not favored for topological reasons.

The family U^IG^Cn ($n = 1-23$), which involves the most stable tautomers of uracil and glycine, was characterized in our recent papers.[34,35] The large number of the cyclic U^IG^C structures results from the fact that U^I possesses four pairs of neighboring proton donor and acceptor sites and G^C possesses six. It was found that the two-body interaction term E_{2-b} correlates with the values of proton affinity and deprotonation enthalpy of the sites involved in hydrogen bonds. It was also demonstrated that formation of a stable cyclic structure with two hydrogen bonds requires not only a favorable two-body interaction E_{2-b} but also a favorable topological match of the proton donor and acceptor sites to minimize the monomer distortion terms E_{1-b}'s. We will use these guidelines when designing complexes formed by minor tautomers of uracil with canonical and zwitterionic glycine.

First we consider which rotamers of U should be tested in complexes with glycine. We notice that the most stable rotamers are topologically better fit to form cyclic hydrogen bonds than the less stable rotamers (see Figures 3 and 4). In addition, isolated higher energy rotamers are less stable than the corresponding most stable rotamer by 3−11 kcal/mol for the monohydroxy tautomers and 1−5 kcal/mol for the dihydroxy tautomer.[6] For these reasons we limited our search to the most stable rotamer of every tautomer of uracil.

The first selection step leaves too many structures for the DFT/B3LYP evaluation. The reason is that every tautomer of uracil possesses more than one pair of neighboring proton donor and acceptor sites capable of forming hydrogen bonds, and G^C possesses six such pairs. For instance, U^{II} possesses three pairs involving electronegative atoms O or N and two extra pairs, in which a CH acts as a proton donor. Six donor−acceptor pairs of G^C multiplied by five complementary acceptor−donor pairs of uracil yields 30 initial structures of the $U^{II}G^C$ complex to be explored. For each U^R, the values of PA and DPE (see Table 2) for proton acceptor and donor sites, respectively, will be used to identify sites capable of forming the strongest hydrogen bonds and thus to further prescreen the set of possible cyclic hydrogen-bonded structures.

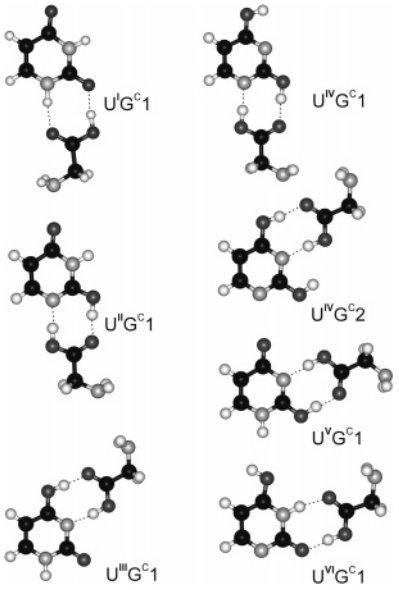


Figure 3. Lowest energy structures of complexes between tautomers of uracil and canonical glycine optimized at the B3LYP/6-31++G** level.

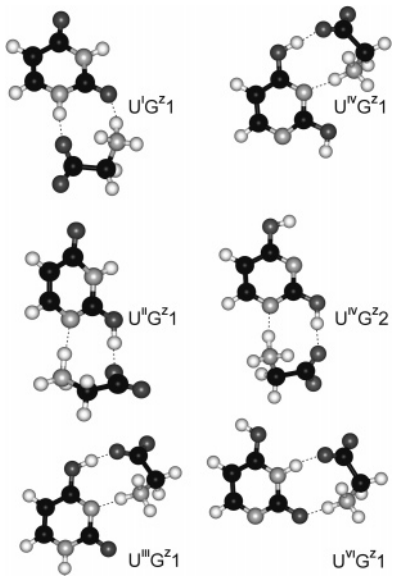


Figure 4. Lowest energy structures of complexes between tautomers of uracil and zwitterionic glycine optimized at the B3LYP/6-31++G** level.

First, we concentrate on the most promising binding sites for minor tautomers of uracil. The simplest are U^{III} and U^V since their sites having the largest value of PA and the smallest value of DPE neighbor each other. These are N3 and O8H in U^{III} and N3 and O7H in U^V. For U^{II}, however, the most acidic site, O7H, and the most basic site, the C5 side of O8, are not first neighbors (see Table 2). Fortunately, the O7H group is located in the proximity of a relatively basic center N1, and these binding sites of U^{II} will be further considered in complexes with glycine. We discard the most basic site, which is the C5 side of O8, since it would have to be paired with the C5H proton donor site, which has a very large value of DPE (see Table 2). Another possibility for U^{II} would be to form hydrogen bonds through N3H and O8. The PA values of N1 and the N3 side of O8 are almost identical, but the DPE value of O7H is ca. 15 kcal/mol

(56) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094.

lower than that for the N3H site. Thus, the O7H and N1 pair remains our top selection for U^{II}.

For U^{VI} the most basic (N1) and acidic (O8H) sites are also spatially separated, and we select N3H and O7 for binding with glycine. Finally, the U^{IV} dihydroxy tautomer has two pairs of neighboring strong proton donor and acceptor sites, i.e., N1 and O7H as well as N3 and O8H (see Table 2). Hence, these two pairs will be further tested in complexes with glycine.

The canonical and zwitterionic tautomers of glycine have different hydrogen-bonding capabilities (see Table 2). In zwitterionic glycine there is only one type of proton donors (NH) and one type of proton acceptors (O9 or O10). The canonical glycine possesses formally six pairs of proton donor and acceptor sites,[35] but two of them have clearly favorable values of PA and DPE. These are the O10H and N13 as well as O10H and O9 pairs. The N13 site is more basic than O9, and the U^IG^C complexes formed by the former pair are characterized by larger values of E_{2-b} than those formed by the latter.[35] Unfortunately, the O10H and N13 pair is topologically not well fit to form cyclic hydrogen bonds with uracil, and the resulting complexes are destabilized by the relatively large values of $E_{1-b}{}^G$.[35] For this reason, in the current study we engage primarily the O9 and O10H sites of canonical glycine in hydrogen bonds with uracil.

Our selection of preferable binding sites for minor tautomers of uracil and canonical and zwitterionic glycine, which is based primarily on the values of PA and DPE as well as on topological arguments, was verified and confirmed at the PM3 level of theory. Moreover, a satisfactory agreement was found between the PM3 and DFT/B3LYP results for selected U^RG^Ln structures. The most stable conformations in the U^RG^Cn and U^RG^Zn families are presented in Figures 3 and 4, respectively. For the U^VG^Z1 complex, a spontaneous double-proton transfer (O7H···O10 → O7···HO10 and H_2N13H···N3 → H_2N13···HN3) was observed at both the B3LYP/6-31++G** and MP2/6-31+G** levels. A product of this transformation is equivalent to U^IG^C19.[35] Thus, a complex between U^V and G^Z is not presented in Figures 4−7.

The geometrical parameters, which characterize hydrogen bonds in the U^RG^Cn and U^RG^Zn structures, are provided in the Supporting Information. The strength of a hydrogen bond is determined by (i) the charge distribution in the proton donor (YH) and acceptor (X), (ii) the distance between H and X, and (iii) the X···HY angle. The two hydrogen bonds involved in the cyclic structures are almost linear. The deviation from linearity amounts at most to 11°. These data suggest that the tautomerization rather than the deformation component will dominate the E_{1-b} terms.

The stabilization energy in the U^IG^C1 complex is about 16 kcal/mol,[35] hence typical for ringlike structures such as the formic acid dimer[57] or the formamide dimer.[58] The values of the X···H distances in other U^RG^C structures are frequently shorter than that in U^IG^C1. The O···HO distance is as short as 1.568, 1.539, and 1.601 Å in $U^{II}G^C1$, U^VG^C1, and $U^{III}G^C1$, respectively, whereas it is 1.663 Å in U^IG^C1. These data suggest that the values of the two-body interaction term E_{2-b} can be larger in complexes of uracil's rare tautomers with G^C than in U^IG^C1. An even more pronounced shortening of the X···H distances takes place in complexes with zwitterionic glycine. The O···H distance is as short as 1.383 and 1.410 Å in $U^{II}G^Z1$ and $U^{III}G^Z1$, respectively. Such distances are usually characteristic for hydrogen bonds involving ions and clearly display the zwitterionic character of glycine.

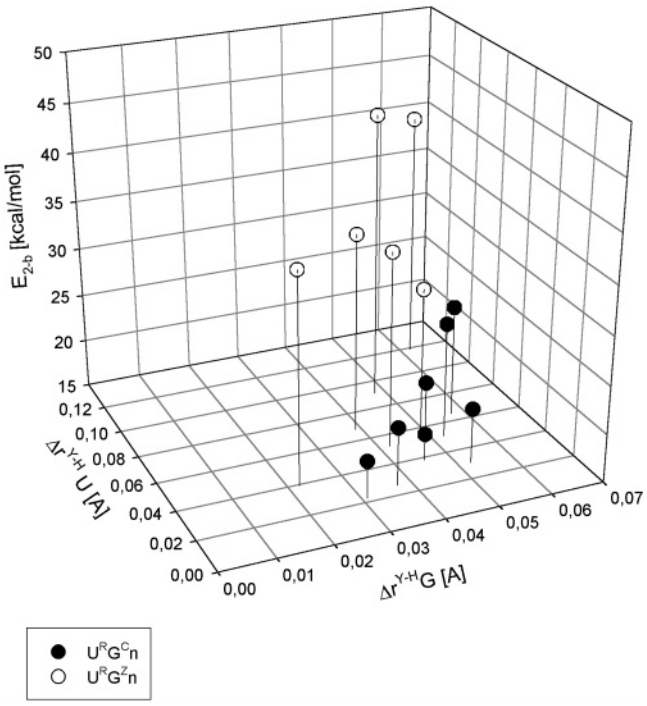


Figure 5. Two-body interaction energy, E_{2-b} (see also eq 5), as a function of the elongation of the proton donor Y−H bonds.

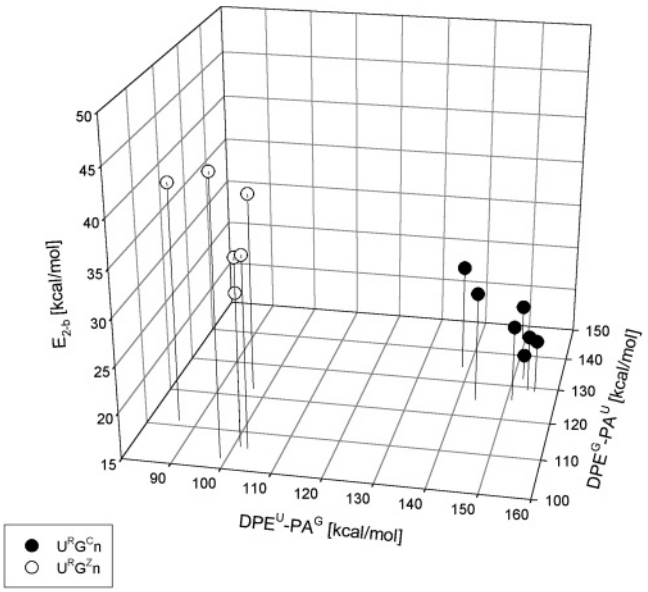


Figure 6. Dependence of the two-body interaction energy E_{2-b} on $DPE^U - PA^G$ and $DPE^G - PA^U$ for the 13 structures depicted in Figures 3 and 4. $DPE^U - PA^G$ and $DPE^G - PA^U$ characterize the propensity of proton donor and acceptor pairs of uracil and glycine to form hydrogen bonds.

C. Energetic Characteristics of the U^RG^Ln Complexes. The B3LYP/6-31++G** values of E_{stab}, H_{stab}, and G_{stab} for the 13 systems are collected in Table 3. In addition, we included the U^IG^Z2 complex, because it provides an avenue for double intermolecular proton transfer, which leads to the U^{III} tautomer (see section III.D). The most stable structure is U^IG^C1. In terms

(57) Pecul, M.; Leszczynski, J.; Sadlej, J. *J. Chem. Phys.* **2000**, *112*, 7930−7938.
(58) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963−4968.

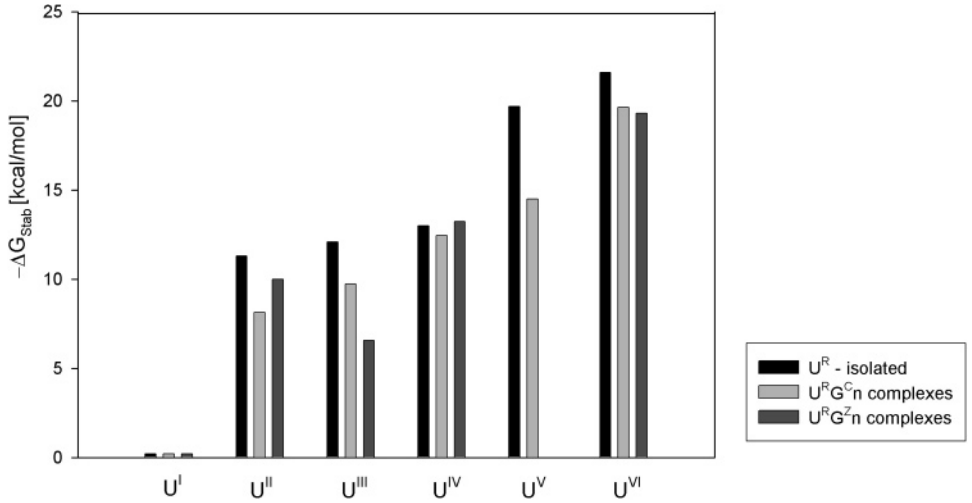


Figure 7. Comparison of the relative free energies (ΔG_{stab}, in kcal/mol) for the U^R, U^RG^Cn, and U^RG^Zn sets of structures.

Table 3. Values of $E_{1-b}{}^{G}$, $E_{1-b}{}^{U-tau}$, $E_{1-b}{}^{U-def}$, E_{2-b}, E_{stab}, H_{stab}, and G_{stab} Calculated at the B3LYP/6-31++G** Level (All Quantities in kcal/mol)

complex	$E_{1-b}{}^{G}$	$E_{1-b}{}^{U-tau}$	$E_{1-b}{}^{U-def}$	E_{2-b}	$E_{stab}{}^{a}$	$H_{stab}{}^{a}$	$G_{stab}{}^{a}$
U^IG^C1	−1.6	0.0	−0.9	18.9	16.4	15.1	4.1
$U^{II}G^C1$	−5.2	−11.1	−2.4	26.6	7.8	7.4	−5.3
$U^{III}G^C1$	−2.6	−11.9	−2.3	23.3	6.4	5.7	−5.7
$U^{IV}G^C1$	−2.5	−12.8	−1.7	20.8	3.8	3.0	−8.5
$U^{IV}G^C2$	−2.0	−12.8	−2.8	17.8	0.3	−0.4	−11.8
U^VG^C1	−3.4	−19.5	−2.9	27.0	1.3	0.9	−10.5
$U^{VI}G^C1$	−2.2	−21.4	−1.3	21.1	−3.8	−4.4	−15.4
U^IG^Z1	−30.1	0.0	−2.1	36.8	4.5	3.0	−9.1
U^IG^Z2[b]	−22.1	0.0	−1.8	23.0	−0.9	−2.1	−13.7
$U^{II}G^Z1$	−27.9	−11.1	−8.3	40.6	−6.8	−6.6	−19.2
$U^{III}G^Z1$	−28.1	−11.9	−7.2	44.5	−2.9	−3.2	−15.8
$U^{IV}G^Z1$	−26.8	−12.8	−5.1	35.6	−9.2	−9.8	−22.4
$U^{IV}G^Z2$	−26.0	−12.8	−4.8	30.4	−13.2	−13.7	−26.2
$U^{VI}G^Z1$	−26.1	−21.4	−3.2	35.2	−15.5	−16.1	−28.3

[a] Positive values of E_{stab}, H_{stab}, and G_{stab} are measures of the strength of hydrogen bonds. [b] The second most stable complex between U^I and G^Z is included because it provides an avenue for double intermolecular proton transfer, which leads to the U^{III} tautomer (see section III.D).

of E_{stab}, it is more stable by 8.6 and 11.9 kcal/mol than the next U^RG^C and the most stable U^RG^Z structure, respectively. Moreover, U^IG^C1 is the only complex, formation of which is favored thermodynamically in the gaseous mixture of uracil and glycine at 298 K; i.e., the value of G_{stab} is positive and amounts to 4.1 kcal/mol. This indicates that over 90% of uracil and glycine would be bound in the U^IG^C1 complex for initial concentrations of U and G of 0.1 mol/dm^3. Other U^RG^C complexes are either stable in terms of H_{stab} by only a few kilocalories per mole or are unstable. It should be reemphasized that in biological systems the H(N1) proton cannot participate in tautomerizations because pyrimidine bases are connected to the sugar unit through the N1 atom.

The stabilization functions E_{stab}, H_{stab}, and G_{stab} for a U^RG^L complex are defined with respect to the most stable tautomers of isolated U and G (see eq 1). Hence, the values of E_{stab} result from a balance between the stabilizing E_{2-b} and destabilizing E_{1-b} components. The E_{1-b} term for glycine contains both the tautomerization and deformation components. Although zwitterionic structure of glycine is not a local minimum, our estimation of $E_{1-b}{}^{G-tau}$ is −23.7 kcal/mol (see section III.A). It is then not surprising that the values of $E_{1-b}{}^{G}$ in the U^RG^Zn complexes are in a range of −26.0 to −30.1 kcal/mol (see Table 3), which implies that the values of the deformation term $E_{1-b}{}^{G-def}$ are in a range −2.3 to −6.4 kcal/mol. The significant values of $E_{1-b}{}^{G-tau}$ and $E_{1-b}{}^{U-tau}$ contribute heavily to the instability of complexes formed by the less stable tautomers. The destabilizing deformation terms, $E_{1-b}{}^{G-def}$ and $E_{1-b}{}^{U-def}$, are also significant despite our best efforts to design topologically fit pairs of U^R and G^L. The values of the deformation term are usually larger in complexes with zwitterionic rather than canonical glycine, suggesting a stronger interaction in the former class of complexes.

The most important result of the current study is that the stabilizing interaction energy term E_{2-b} can be much larger in complexes formed between the rare tautomers of U and G than in the U^IG^C1 complex, thus compensating, at least partially, for the unfavorable $E_{1-b}{}^{X-tau}$ terms (X = U or G). This enhancement of the E_{2-b} term is typically larger in complexes with zwitterionic glycine and is as large as 25.6 and 21.7 kcal/mol for the $U^{III}G^Z1$ and $U^{II}G^Z1$ complexes, respectively. It is smaller in complexes with canonical glycine but is still as large as 8.1, 7.7, and 4.4 kcal/mol for U^VG^C1, $U^{II}G^C1$, and $U^{III}G^C1$, respectively. A significant stabilization of U^{III} is of particular interest because this tautomer of uracil and thymine has binding sites matching those of guanine rather than those of adenine (see Figure 1).

The cyclic structures presented in Figures 3 and 4 are stabilized by two hydrogen bonds. Formation of a hydrogen bond X···HY is usually accompanied by an elongation of the Y−H bond, Δr^{Y-H}, and a red shift of the frequency of the Y−H stretching mode. In Figure 5, we display the two-body interaction energy (E_{2-b}) as a function of two variables: the elongation of a Y−H bond in uracil (Δr^{Y-H} U) and glycine (Δr^{Y-H} G).

A Y−H bond of uracil in the U^RG^Zn complexes interacts with a very basic −COO− group of the glycine zwitterion (the PA of this group exceeds by 9.3 kcal/mol that of the second most basic site, which occurs at N1 of U^{VI}), whereas in the U^RG^Cn complexes it interacts with a less basic carbonyl oxygen of the COOH group. Consequently, the values of Δr^{Y-H} U are usually larger in complexes with the zwitterionic than with canonical glycine (see Figure 5). The values of Δr^{Y-H} G cover a similar range for both tautomeric forms of glycine. The largest values

of Δr^{Y-H} are 0.12 and 0.06 Å for U (an O−H bond) and G (an N−H bond), respectively.

The perturbations of covalent Y−H bonds correlate with the magnitude of the E_{2-b} term and provide an insight into the strength of individual hydrogen bonds. The 3-D plot in Figure 5 shows that large values of E_{2-b} occur only when *both* hydrogen bonds involve significantly perturbed proton donors. It is worth noting, nevertheless, that an extra interaction can develop which contributes to the stability of a complex. The pair $U^{II}G^{Z}1$ and $U^{III}G^{Z}1$ provides an excellent example. The $U^{II}G^{Z}1$ complex, with $E_{2-b} = 40.6$ kcal/mol, is characterized by the largest values of Δr^{Y-H} G and Δr^{Y-H} U. However, the value of E_{2-b} for the $U^{III}G^{Z}1$ complex is even larger and amounts to 44.5 kcal/mol. Apparently, an extra hydrogen bond in $U^{III}G^{Z}1$ between the N13−H site of glycine and the CO7 carbonyl of uracil provides an additional stabilization, although relatively weak, as the H···O7 distance is 2.87 Å.

A plot of E_{2-b} versus *x1* and *x2* is presented in Figure 6 for the 13 $U^{R}G^{L}n$ complexes. The largest values of E_{2-b} occur for small values of both *x1* and *x2*, with the results being more sensitive to the latter variable. This is again a manifestation of the large PA for the −COO− group of zwitterionic glycine. The data displayed in Figure 6 demonstrate, however, the existence of some irregularities, which may result from additional hydrogen bonds and/or factors other than the values DPE and PA which could also contribute to the strength of a cyclic hydrogen bond. Nevertheless, we tried to establish a quantitative relation between *x1*, *x2*, and E_{2-b}. The first-order polynomial expansion of E_{2-b} in terms of *x1* and *x2* gives a relatively small value of $r^2 = 0.874$. No significant improvement is accomplished with a functional form $E_{2-b} = ax1^2 + b*x2^2 + c$. Finally, the value of r^2 increases to 0.899 for a functional $E_{2-b} = ax1^2 + bx1 + cx2^2 + dx2 + ex1x2 + f$. In the latter case, however, a larger value of r^2 may result from a bigger number of regression parameters. Thus, the quality of a fit based on *x1* and *x2* only, even with a multiparameter functional form, is relatively low, which demonstrates inherent complexity of the hydrogen-bonding phenomenon. Simple characteristics of monomers, like DPE and PA of centers directly involved in the interaction, are not able to account for subtle effects accompanying formation of hydrogen bonds, such as cooperativity in cyclic systems, valence repulsion, or polarization interactions.

One theory of spontaneous mutations in DNA/RNA assumes formation of a mutationally significant number of minor tautomers of NBs as a consequence of the thermodynamic equilibrium that attains in physiological conditions.[1] The previous genetic experiments, regarding proofreading activity of DNA polymerases,[59] demonstrated that the frequency of spontaneous mutations falls in a range of $10^{-6}-10^{-9}$. For isolated uracil, the instability of the three most stable rare tautomers on a free enthalpy scale ($T = 298$ K) is as large as 11−13 kcal/mol, which implies that a rare tautomer appears once in $10^{8}-10^{10}$ bases. Our results presented in Table 3 suggest, however, that interactions of nucleic bases with potent molecules, having several proton-donor/acceptor centers, like an amino acid, may dramatically enhance the population of rare tautomers.

As we pointed out, the large two-body interaction in complexes of glycine with rare tautomers of uracil may favor a rare tautomer over the most stable one. According to our data, formation of $U^{I}G^{C}1$ in the gaseous mixture of uracil and glycine is favored thermodynamically (Table 3). Depending on the relative free enthalpies of the remaining species, a small fraction of the $U^{I}G^{C}1$ complex will convert to other $U^{R}G^{L}n$ complexes in the equilibrated gaseous mixture. In Figure 7, we compare relative free enthalpies of the isolated tautomers of uracil (black bars) with those of the tautomers interacting with the canonical (gray bars) as well as with zwitterionic glycine (dark gray bars). The free enthalpies of U^{I}, $U^{I}G^{C}1$, and $U^{I}G^{Z}1$ have been assumed as reference points for the families U^{R}, $U^{R}G^{C}n$, and $U^{R}G^{Z}n$, respectively. One should keep in mind, however, that in the gas phase $U^{I}G^{Z}1$ is less stable than $U^{I}G^{C}1$ by 13.2 kcal/mol in terms of G_{stab}. This probably will not be the case in water solution, where the zwitterionic form of glycine is more stable than the canonical.

The interaction with glycine increases the relative stability of rare tautomers of uracil (see Figure 7). In the case of complexes with canonical glycine, the ordering of rare tautomers remains the same as for isolated uracil. However, the ordering is changed in complexes with zwitterionic glycine, with $U^{III}G^{Z}1$ becoming the second most stable complex. For this complex, the enhancement of the relative stability is the largest and amounts to 5.2 kcal/mol.

Both $U^{II}G^{C}1$ and $U^{III}G^{C}1$ are unstable with respect to $U^{I}G^{C}1$ by less than 10 kcal/mol (see Figure 7). This increased stability of rare tautomers is reflected in the population ratios at $T = 298$ K: $[U^{II}G^{C}1]:[U^{I}G^{C}1] = 7 \times 10^{-7}$ and $[U^{III}G^{C}1]:[U^{I}G^{C}1] = 6 \times 10^{-8}$, which fall in the experimentally determined range of substitution frequency observed through the proofreading activity of DNA polymerases.[59] The relative stability of rare tautomers is even further enhanced in the $U^{R}G^{Z}n$ family. The $U^{III}G^{Z}1$ structure is unstable with respect to $U^{I}G^{Z}1$ by 6.7 kcal/mol, which leads to a population ratio $[U^{III}G^{Z}1]:[U^{I}G^{Z}1] = 1 \times 10^{-5}$. A population of 1×10^{-5} would indicate the possibility of spontaneous mutations as a result of tautomeric equilibrium. However, one should keep in mind that the overall population of complexes with zwitterionic glycine is very low in the gas phase. Hence, a critical issue for our future research is to determine the relative stability of the $U^{R}G^{Z}$ structures in water solution, in which zwitterionic glycine is the dominant tautomer. The increased stability of the type-III tautomer of uracil, and probably thymine, can be relevant for point mutations because tautomers of this type can be involved in mispairing with guanine (see Figure 1).

D. Glycine-Assisted $U^{I}\leftrightarrow U^{III}$ Tautomerization. The structures of stationary points for three possible $U^{I}\leftrightarrow U^{III}$ processes are displayed in Figure 8. The upper graph demonstrates an intramolecular proton transfer for isolated uracil. The B3LYP/6-31++G** reaction barriers for the forward and reverse tautomerizations are 44 and 32 kcal/mol, respectively. These barriers are much too high to enable attaining the equilibrium concentrations within the reproduction period for typical biological species, as the value of $\tau_{99.9\%}$ amounts to 1.22×10^{11} s, i.e., 3876 years!

Much smaller barriers are predicted for glycine-assisted tautomerizations (see Figure 8, middle and bottom). For $U^{I}G^{C}\leftrightarrow U^{III}G^{C}$, the barriers of 10.8 and 2.5 kcal/mol are predicted for the forward and reverse processes, respectively. Here, 99.9% of the equilibrium concentration of U^{III} is reached within 2.8 ×

(59) Fersht, A. R.; Knill-Jones, J. W. *J. Mol. Biol.* **1983**, *165*, 633−634.

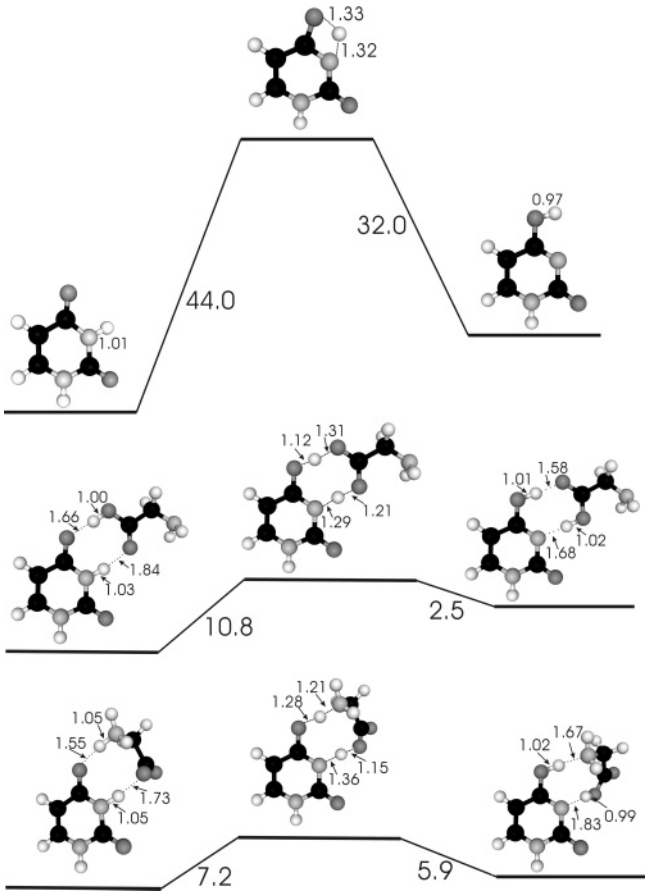


Figure 8. Structures of stationary points for proton-transfer reactions leading to the U^{III} tautomer. Intramolecular tautomerization $U^{I} \leftrightarrow U^{III}$ (top); tautomerization assisted by canonical glycine, $U^{I}G^{C} \leftrightarrow U^{III}G^{C}$ (middle); and tautomerization assisted by zwitterionic glycine, $U^{I}G^{Z} \leftrightarrow U^{III}G^{Z}$ (bottom). The barriers (differences in electronic energies) for the forward and reverse reactions are displayed in kilocalories per mole. The dotted lines indicate hydrogen bonds, while continuous lines show intermediate "bonds" at transition states. The bond lengths are given in angstroms.

10^{-11} s, which is many orders of magnitude smaller than a reproduction period. However, the 2.5 kcal/mol barrier for the reverse reaction leads to a half-life of 2.8×10^{-12} s, and tunneling effects will further facilitate the reverse process. Hence, $U^{III}G^{C}$ produced in the double-proton-transfer process represents an unstable reaction intermediate, which is quickly transformed back into $U^{I}G^{C}$ in the time scale of the nucleotide−amino acid interaction. However, after dissociation of the glycine from $U^{III}G^{C}$, the U^{III} tautomer would be a long-lived species, as the barrier for the transformation back to U^{I} is 32 kcal/mol (see the top of Figure 8).

A complex $U^{I}G^{Z}2$, which meets both the thermodynamic and kinetic requirements for a spontaneous DNA mutation, is characterized in Table 3 and in the bottom of Figure 8. The forward and reverse barriers for $U^{I}G^{Z} \leftrightarrow U^{III}G^{C}$ are 7.2 and 5.9 kcal/mol, respectively. Here, $\tau_{99.9\%}$ amounts to 7.8×10^{-9} s, which is still many orders of magnitude smaller than a reproduction period. In addition, the 5.9 kcal/mol barrier for the reverse reaction leads to a half-life of 9×10^{-10} s. Finally, the $U^{III}G^{C}$ structure is unstable with respect to $U^{I}G^{Z}$ by only 1.3 kcal/mol, which leads to a population ratio $[U^{III}G^{C}]:[U^{I}G^{Z}] = 1 \times 10^{-1}$. To conclude, the $U^{I} \rightarrow U^{III}$ tautomerization is kinetically and thermodynamically forbidden in isolated uracil, whereas it is allowed when assisted by glycine, in particular by its zwitterionic form.

IV. Summary

Six tautomers of uracil and their complexes with canonical and zwitterionic glycine have been studied at the density functional B3LYP/6-31++G** level of theory. The significant thermodynamic instability of rare tautomers of uracil, treated as isolated species, does not support a mechanism of spontaneous mutations, which is based on thermodynamic equilibrium between various tautomers.

Several complexes between six tautomers of uracil and canonical and zwitterionic glycine have been designed on the basis of (a) proton affinity and deprotonation enthalpy of proton acceptor and donor binding sites, respectively, and (b) topological fitness of binding sites required to form cyclic hydrogen-bonded structures free of a significant internal strain. Initial explorations, performed at the PM3 level of theory, were followed by the B3LYP/6-31++G** geometry optimizations and frequency calculations. Thirteen $U^{R}G^{L}$ structures (U^{R} and G^{L} denote the *R*th and *L*th tautomers of uracil and glycine, respectively) stabilized by two almost linear hydrogen bonds have been characterized. A complex formed by the most stable tautomers is stable in terms of Gibbs free energy by 4.1 kcal/mol. Other complexes are unstable with respect to the lowest tautomers of isolated U and G moieties. The stabilization energies (E_{stab}) for $U^{R}G^{L}$'s were decomposed into the destabilizing one-body (E_{1-b}) terms and stabilizing two-body (E_{2-b}) interaction energies. The E_{1-b} terms are related to the energy of (a) tautomerization of the unperturbed moieties and (b) distortions of the resulting rare tautomers in the complex. The two-body term is related to the interaction energy between distorted tautomers. An important finding from the current study is that the stabilizing interaction energy term E_{2-b} can be much larger in binary complexes of rare tautomers of U and G than in the complex formed by the most stable tautomers. The large values of E_{2-b} compensate, at least partially, for the unfavorable E_{1-b} terms. A significant enhancement of the E_{2-b} terms has been observed in complexes formed by the U^{III} (2-oxo-4-hydroxy) and U^{II} (2-hydroxy-4-oxo) tautomers of uracil.

Formation of a hydrogen bond X···HY is usually accompanied by an elongation of the Y−H bond, Δr^{Y-H}. We found that the values of Δr^{Y-H} in the $U^{R}G^{L}$ complexes correlate with the magnitude of the E_{2-b} term. The largest values of Δr^{Y-H} were found to be 0.07 and 0.12 Å in complexes with the canonical and zwitterionic glycine, respectively, with the corresponding values of E_{2-b} being 26.6 and 40.6 kcal/mol. The values of E_{2-b} also correlate with the values of deprotonation enthalpy and proton affinity of the monomer sites involved in hydrogen bonds.

The interaction with glycine increases the relative stability of rare tautomers of uracil. The ordering of tautomers is the same for isolated uracil and for its complexes with canonical glycine, but the magnitude of instability is suppressed to 9.4 and 9.8 kcal/mol for U^{II} and U^{III}, respectively. The increased stability of rare tautomers is reflected in the population ratios at $T = 298$ K: $[U^{II}G^{C}1]:[U^{I}G^{C}1] = 7 \times 10^{-7}$ and $[U^{III}G^{C}1]:[U^{I}G^{C}1] = 6 \times 10^{-8}$.

The relative stability of rare tautomers is even further enhanced in complexes with zwitterionic glycine. Moreover,

the ordering of tautomers is switched when compared with that of the isolated uracil. The $U^{III}G^{Z}1$ complex becomes the second most stable, with $U^{I}G^{Z}1$ being more stable by only 6.7 kcal/mol, which leads to a population ratio $[U^{III}G^{Z}1]:[U^{I}G^{Z}1] = 1 \times 10^{-5}$. A population of 1×10^{-5} would indicate a possibility of spontaneous mutations as a result of the tautomeric equilibrium.

The kinetic barriers for the conversion of U^{I} into U^{III} are much larger for isolated uracil than for uracil bound to glycine. The barriers for the forward process are 44.0, 10.8, and 7.2 kcal/mol for $U^{I} \leftrightarrow U^{III}$, $U^{I}G^{C} \leftrightarrow U^{III}G^{C}$, and $U^{I}G^{Z} \leftrightarrow U^{III}G^{C}$, respectively. The corresponding barriers for the reverse process are 32.0, 2.5, and 5.9 kcal/mol. The barriers for the $U^{I}G^{C} \leftrightarrow U^{III}G^{C}$ and $U^{I}G^{Z} \leftrightarrow U^{III}G^{C}$ are sufficiently small to make the lifetime of the canonical base smaller than the lifetime of biological species. The $U^{I}G^{Z} \leftrightarrow U^{III}G^{C}$ process is particularly interesting as the accessible barrier for the forward reactions is accompanied by a sufficiently high barrier of the reverse reaction to make the process relevant to spontaneous mutations.

The actually operating mechanism of any biochemical process cannot be proved through gas-phase ab initio calculations. The calculations of this type can, however, identify probable mechanisms that require further experimental and computational (multiscale) verifications. The increased stability reported here of the type-III tautomer of uracil, and probably thymine, in complexes with canonical and zwitterionic glycine can be relevant for point mutations because its proton donor and acceptor sites are complementary with those of guanine rather than those of adenine (see Figure 1). The relative stability of the $U^{R}G^{Z}$ structures in water solution and of complexes of thymine with glycine will be discussed in our future reports.

Acknowledgment. We thank a reviewer for a suggestion to study glycine-assisted proton transfer in the uracil−glycine complex. J.R. and I.D. were supported by the Polish State Committee for Scientific Research (KBN), Grant No. 4 T09A 012 24. M.G. was supported by the U.S. DOE, Office of Biological and Environmental Research, Low Dose Radiation Research Program. His research was performed in the Pacific Northwest National Laboratory (PNNL). Battelle operates the Pacific Northwest National Laboratory for the Department of Energy. The computer time allocations provided by the Academic Computer Center in Gdańsk (TASK) is gratefully acknowledged. This research also used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Supporting Information Available: Selected geometrical characteristics of hydrogen bonds in complexes between minor tautomers of uracil and glycine. This material is available free of charge via the Internet at http://pubs.acs.org.

JA048730K