Polish J. Chem., **76**, 1243–1247 (2002)

On the Stability of Uracil–Glycine Hydrogen-Bonded Complexes. A Computational Study*

by I. Dąbkowska[1], M. Gutowski[2,1] and J. Rak[1]**

[1]*Department of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland*
[2]*Environmental Molecular Sciences Laboratory, Theory, Modeling&Simulation, Pacific Northwest National Laboratory, Richland, WA 99352, USA*

(Received December 21st, 2001; revised manuscript February 13th, 2002)

The topological space of the glycine–uracil complex has been scanned at the semi-empirical PM3 level of theory and the lowest energy complexes have been characterized at the B3LYP/6-31++G** level. These complexes are characterized by two hydrogen bonds, in which the carboxylic group of glycine interacts with proton donor and acceptor sites of uracil. The stabilization energy for three complexes, with the NH···O and O···HO hydrogen bonds each, spans a narrow range of 15.6–12.3 kcal/mol. The fourth complex with one strong O···HO and one weak CH···O hydrogen bond is bound by 10.2 kcal/mol.

Key words: uracil–glycine complexes, hydrogen bond, PM3 and DFT calculations

The binding of proteins to DNA plays an important role in the regulation and control of gene expression. Hydrogen bonds between the hydrophilic side chain of a peptide bond or an amino acid and DNA are among the most important interactions in nature as they are responsible for high specificity of protein binding [1]. Therefore, basic knowledge concerning the interactions between the building blocks of proteins and DNA – amino acids and nucleic acid bases – is of great interest. Although these systems are definitely simpler than real biochemical targets, quantitative information regarding the interactions between amino acid – nucleobase pairs can, in principle, be used to generate high quality potentials, which can be applied in modeling of macromolecules.

The interactions between amino acid side chains of amino acids and nucleobases were investigated theoretically, by using the density functional theory (DFT) method [2], as well as experimentally by the field ionization mass spectrometry technique [3]. The Rydberg electron transfer experiments on dipole-bound anions of adenine bound to imidazole, pyrole and methanol (models for serine and threonine) provided some information about the neutral complexes [1]. In this paper we report on the results of semiempirical PM3 [4] and DFT calculations concerning the simplest amino acid – nucleobase complex, *i.e.*, the dimer of glycine and uracil. Glycine is the simplest amino acid and uracil is a building block of RNA, which is analogous to thymine pres-

---

* Presented at the 1st Russian-Ukrainian-Polish Conference on Molecular Interactions in Gdańsk, June 2001, (Poland).

** To whom correspondence should be addressed.

ent in DNA. The main goal of our efforts is to recognize the decisive factors responsible for the stability of the uracil–glycine complexes in the gas phase. These findings may help to interpret the recently measured photoelectron spectrum of the anion of the uracil–glycine complex [5].

COMPUTATIONAL METHODS

We applied both the semiempirical PM3 methods [4] as well as the DFT method with a hybrid B3LYP functional [6–8] and 6-31++G** basis sets [9,10] to study the structure and stability of the title complexes. Our recent work on the neutral and charged arginine and its dipole-bound anion [11,12], as well as other reports on complexes between nucleic acid bases and water [13–15], demonstrated usefulness of the B3LYP method with 6-31G** basis set supplemented with diffuse functions in studying hydrogen-bonded systems. The low-energy structures were initially identified at the PM3 level of theory. This method was used for a preliminary search of the potential energy surface to select candidate structuress for more accurate investigations. Next, full geometry optimizations and frequency calculations have been performed at the B3LYP level. All calculations were carried out with the MOPAC2000 [16] and GAUSSIAN 98 [17] codes.

Both glycine and uracil belong to the class of molecules having several proton donor and acceptor centers capable of forming hydrogen bonds of various strengths. These are O(7), O(8), N(1), N(3), and C(5) for uracil and N(13), O(9), O(10) for glycine; see Fig. 1. Here we focused on complexes with *two* intermolecular hydrogen bonds, as topological reasons disable formation of systems, where three strong hydrogen bonds are present. Therefore, the global minimum is expected to be a complex with two strong hydrogen bonds. The five proton donor or acceptor sites present in uracil create four regions (see UG1 in Fig. 1) capable of forming two hydrogen bonds at the same time, with one site acting as a proton donor and another as a proton acceptor. On the other hand, glycine possesses one proton-acceptor (O(9)) and two proton donor-and-acceptor centers (O(10) and N(13)), which can be assembled in six proton accep-

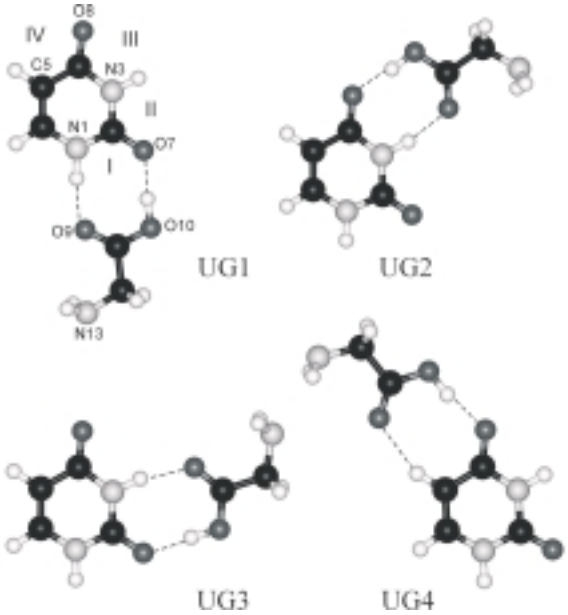

Figure 1. B3LYP/6-31++G** optimized geometries of the UG1–UG4 complexes. I, II, III, and IV denote the regions in uracil capable of forming two adjacent hydrogen bonds.

tor-donor pairs. Six proton-acceptor pairs of glycine times four complementary proton-acceptor pairs of uracil yields twenty four complexes that were inspected at the PM3 level of theory.

The geometry optimizations carried out at the B3LYP/6-31++G** level proved that the most stable complexes are those with the proton donor and acceptor centers of uracil, interacting with the carbonyl (O(9)) and hydroxyl (O(10)) oxygens of glycine. Four possible structures of this type are shown in Fig. 1. We will restrict the following discussion to the complexes displayed in this figure. These are, however, the most stable structures, and therefore, they should predominate in the equilibrated gas phase mixture of uracil and glycine.

RESULTS AND DISCUSSION

In Table 1 we summarize the relative stabilities of the four most stable uracil–glycine (UG) complexes. The UG1–UG4 structures have been ordered according to their decreasing stability. The stability of UGn is expressed in terms of E_{stab} and H_{stab}. E_{stab} is defined as a difference in electronic energies of the monomers and the dimer

$$E_{stab} = E^{U}(Geom^{U}) + E^{G}(Geom^{G}) - E^{UG}(Geom^{UG})$$

with the electronic energy E^{X} (X = U, G, UG) computed for the coordinates determining the optimal geometry of X (*i.e.*, the geometry where E^{X} is at the minimum). E_{stab} is decomposed as [18]

$$E_{stab} = E^{U}_{dist} + E^{G}_{dist} + E^{UG}_{\text{int}}$$

where E^{X}_{dist} is a repulsive one-body component related to a distortion of the monomer X (X = U, G) in the dimer

$$E^{X}_{dist} = E^{X}(Geom^{X}) - E^{X}(Geom^{UG})$$

and E^{UG}_{int} is a two-body interaction energy between the distorted monomers [19]

$$E^{UG}_{\text{int}} = E^{U}(Geom^{UG}) + E^{G}(Geom^{UG}) - E^{UG}(Geom^{UG})$$

The E^{UG}_{int} component was corrected for basis set superposition error using the counterpoise method of Boys and Bernardi [20] and the E^{X}_{dist} terms were calculated with monomer centered basis sets [21]. Finally, 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. The values of H_{stab} discussed below were obtained for T = 298 K and p = 1 atm.

Table 1. The values of E^{U}_{dist}, E^{G}_{dist}, E^{UG}_{int}, E_{stab}, and H_{stab} calculated at the B3LYP/6-31++G** level. All quantities in kcal/mol.

Structure	E^{U}_{dist}	E^{G}_{dist}	E^{UG}_{int}	E_{stab}	H_{stab}
UG1	–0.9	–1.6	18.1	15.6	14.3
UG2	–0.8	–1.3	15.4	13.3	12.0
UG3	–0.7	–1.0	14.0	12.3	10.9
UG4	–0.3	–0.8	11.3	10.2	8.7

The stabilization energy varies between 15.6 and 10.2 kcal/mol (see Table 1), which places the uracil–glycine complex among medium bound complexes. The UG1–UG3 complexes, with the NH···O and O···HO hydrogen bonds each, span a narrow range of E_{stab} of 15.6–12.3 kcal/mol. The UG4 complex with one strong O···HO and one weak CH···O hydrogen bond is still relatively strongly bound as E_{stab} and H_{stab} amount to 10.2 and 8.7 kcal/mol, respectively. A striking parallelism between E_{stab} and H_{stab} emphasizes the basic role of two hydrogen bonds in every complex and indicates that the contributions to H_{stab} arising from rotations and vibrations are of secondary importance. The monomer distortion terms follow a clear trend: the stronger is the interaction energy E^{UG}_{int}, the larger (more repulsive) are the monomer distortion terms E^{X}_{dist}. The sum of the monomer distortion terms spans a narrow range of 1.4 kcal/mol (2.5 kcal/mol for the most stable UG1 and 1.1 kcal/mol for the least stable UG4).

The geometrical features of intermolecular hydrogen bonds that are present in the UG1–UG4 structures are summarized in Table 2. The strength of a hydrogen bond is determined by the (i) charge distribution in the proton donor (YH) and acceptor (X) fragment, (ii) the distance between H and X, (iii) and the X····HY angle. As demonstrated by the data gathered in Table 2 and Fig. 1, every structure is stabilized by two hydrogen bonds that differ in length and angle. Judging on the basis of geometrical parameters only, the bond formed between O(10)H of glycine and one of the oxygens (O(7) or O(8)) of uracil is apparently stronger. The H···X distance for the stronger bond is between 1.66 and 1.78 Å and the X···HY angle is very close to 180 degrees (see Table 1). In the second hydrogen bond, the X···H distance and X···HY angle are in the range of 1.78–1.86 Å and 165–171 degrees, respectively. It is worth to note that these geometrical features correlate with the relative stability of UG1–UG4, see Table 1. The shorter and more linear are the hydrogen bonds, the more stable is the complex.

Thus, we demonstrated that the most stable complexes between uracil and glycine are formed when the carboxylic group of glycine is bound through two hydrogen bonds to uracil. The largest stabilization energy of 15.6 kcal/mol, determined at the B3LYP/6-31++G** level, clearly shows that the complex is relatively strongly bound. The work on complexes formed by higher energy tautomers of glycine and uracil is in progress.

Table 2. Selected geometrical characteristics of hydrogen bonds in the uracil–glycine complexes optimized at the B3LYP/6-31++G** level.

Structure[a]	Hydrogen bond type	X···H[b] distance (Å)	X···HY[c] valence angle (degrees)
UG1	O7···HO10	1.663	177.38
	N1H···O9	1.783	170.89
UG2	O8···HO10	1.666	177.22
	N3H···O9	1.836	167.90
UG3	O7···HO10	1.697	177.88
	N3H···O9	1.863	166.12
UG4	O8···HO10	1.729	171.39
	C5H···O9	2.236	154.41

[a]For atomic labels see Fig. 1. [b]X denotes proton acceptor (N or O). [c]X, Y denote electronegative atoms involved in the hydrogen bond.

Acknowledgments

This work was supported by the Polish State Committee for Scientific Research, Grant No. DS/8221-4-0140-2. M.G. was supported by the US DOE, Office of Basic Energy Sciences, Chemical Sciences Division. The computer time allocations provided by the Academic Computer Center in Gdańsk (TASK) is also gratefully acknowledged.

REFERENCES

1. Desfrancois C., Carles S. and Schermann J.P., *Chem. Rev.*, **100**, 3943 (2000).
2. Pichierri F., Aida A., Gromiha M.M. and Sarai A., *J. Am. Chem. Soc*., **121**, 6152 (1999).
3. Galetich I., Kosevich M., Shelkovsky V., Stepanian S.G., Blagoi Y.P. and Adamowicz L., *J. Mol. Struct*., **478**, 155 (1999).
4. Stewart J.J.P., *J. Comput. Chem*., **10**, 209; 221 (1989).
5. Bowen K., unpublished results.
6. Becke A.D., *Phys. Rev. A*, **38**, 3098 (1988).
7. Becke A.D., *J. Chem. Phys*., **98**, 5648 (1993).
8. Lee C., Yang W. and Paar R.G., *Phys. Rev. B*, **37**, 785 (1988).
9. Ditchfield R., Hehre W.J. and Pople J.A., *J. Chem. Phys*., **54**, 724 (1971).
10. Hehre W.J., Ditchfield R. and Pople J.A., *J. Chem. Phys*., **56**, 2257 (1972).
11. Rak J., Skurski P., Simons J. and Gutowski M., *J. Am. Chem. Soc*., **123**, 11695 (2001).
12. Skurski P., Rak J., Simons J. and Gutowski M., *J. Am. Chem. Soc*., **123**, 11073 (2001).
13. Kryachko E.S., Nguyen M.T. and Zeegers-Huyskens T., *J. Phys. Chem. A*, **105**, 1934 (2001).
14. Chandra A.K., Nguyen M.T. and Zeegers-Huyskens T., *J. Phys. Chem. A*, **102**, 6010 (1998).
15. Chandra A.K., Nguyen M.T., Uchimaru T. and Zeegers-Huyskens T., *J. Phys. Chem. A*, **103**, 8853 (1999).
16. Stewart J.J.P., MOPAC2000, Fujitsu Limited, (1999).
17. Frisch M.J., *et al*., Gaussian, Inc., Pittsburgh PA, (1998).
18. Kendall R.A., Simons J., Gutowski M. and Chałasiński G., *J. Phys. Chem*., **93**, 621 (1989).
19. Szalewicz K. and Jeziorski B., *J. Chem. Phys*., **109**, 1198 (1998).
20. Boys S.F. and Bernardi F., *Mol. Phys*., **19**, 553 (1970).
21. Chałasiński G. and Gutowski M., *Chem. Rev.*, **88**, 943 (1988).