Polish J. Chem., **75**, 1039–1049 (2001)

A Theoretical Study of the *cis-syn* Cytosine-containing Pyrimidine Dimers in the Gase Phase and a Water Cluster and a Tautomer-Bypass Mechanism for the Origin of UV-induced Mutations

by V.I. Danilov[1], A. Leś[2*] and J.L. Alderfer[3]

[1]*Institute of Molecular Biology and Genetics, The National Academy of Sciences of Ukraine, 150 Zabolotny St., Kiev-143, 03143, Ukraine*
[2]*Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland*
[3]*Roswell Park Cancer Institute, Buffalo, N.Y. 14263 U.S.A.*

(Received February 5th, 2001)

A quantum mechanical study of *cis-syn* cyclobutane cytosine-containing photodimers including the normal and rare tautomeric forms of bases has been performed using the *ab initio* method at HF/6-31G(d,p), MP2(fc)//HF/6-31G(d,p) and MP2(fc)/6-31G(d,p) levels. It is predicted that in the gas phase all photodimers containing the rare imino form of cytosine are more stable than those containing its normal form. The Monte Carlo study of the hydration for cytosine-consisting dimers showed that the dimer containing the imino form of cytosine is stabilized by water cluster more than that containing its amino forms. As a result, the imino form of cytosine in the cytosine-containing dimer directs the incorporation of adenine in the complementary strand during replicative bypass. Data obtained point to the cytosine tautomerism as a possible mechanism for the origin of UV-induced mutation.

Key words: photodimers, tautomerism, UV-mutagenesis, Hartree-Fock, MP2, Monte Carlo calculations

The formation of pyrimidine dimers Pyr[ ]Pyr is the most frequent type of damage induced in DNA by the UV portion of sunlight (for review see [1]). Pyr[ ]Pyr are formed between adjacent pyrimidines on the same strand of DNA by the formation of a cyclobutane ring resulting from saturation of the double bonds in their ring structure. The cyclobutyl type dimers were isolated experimentally from irradiated of native DNA and frozen Thy or Ura solutions.

The fusion of the pyrimidine bases by C(5)–C(6) bonds can give rise to the formation of four types of stereoisomers (see data survey [2]). Thy[ ]Thy cyclobutane dimers are formed most readily, followed by Thy[ ]Cyt or Cyt[ ]Thy; Cyt[ ]Cyt dimers are least abundant. The presence of the pyrimidine dimers will distort the double helical structure of DNA although experiments indicate that the DNA remains double-stranded. The extent of such a backbone distortion caused by pyrimidine dimers

*Corresponding author: e-mail: ales@tiger.chem.uw.edu.pl; phone: (48-22)8222309; fax: (48-22) 8225996.

is determined by the sides of cyclobutane ring. Thus, the two adjacent pyrimidines must be pulled much closer to each other than in normal DNA.

If DNA contains a pyrimidine dimer that cannot be repaired, replication and transcription are blocked at this site. As a result, the UV irradiation of bacteria, phages, and animal cells frequently causes lethal effect. At the same time UV radiation is highly mutagenic due to the formation in nucleic acids of pyrimidine dimers containing cytosine (for review see [1]). UV-induced mutations are caused by point mutations. In a series of studies (*e.g.* see [3–5] and references therein), it has been shown that the majority of point mutations are determined by guanine-cytosine→adenine-thymine (G–C→A–T) or cytosine→thymine (uracil) (C→T (U)) transitions at the dipyrimidine sites and, to a lesser extent, CC→TT double transitions in homopyrimidine sequences, and T→C transitions at dipyrimidine sites. It was detected that the most of mutations occured at the 5′-TpC-3′ and 5′-CpC-3′ sequences and practically none at 5′-TpT-3′ or 5′-CpT-3′ sequences, in both bacteria and human cells [6,7].

The above-mentioned transitions are implicated in certain hereditary cancers. First of all it concerns of nonmelanoma skin cancer in humans. The cyclobutane pyrimidine dimers have been implicated as premutagenic lesions in the initiation of basal and squamous cell carcinomas. Sites of the pyrimidine dimers formation correlate with sites of C→T transition mutations in one of the most frequently mutated genes in human cancers, the p53 tumor suppressor gene in human skin tumors, and the C→T and tandem CC→TT transition mutations are considered the signatures of sunlight-induced carcinogenesis (*e.g.* see [8]). Therefore, UV induced formation of pyrimidine photodimers within DNA is a major cause of the mutagenic and carcinogenic effects of solar light.

The mechanisms by which UV irradiation can generate mutations are not completely understood. In order to account for the predominance of G–C→A–T transition mutations in DNA there was proposed a mechanism conditioned by the photolytic deamination of Cyt in the Cyt-containing pyrimidine dimers (*e.g.* see [9]). The deamination proceeds by the hydrolysis of the imido amide group of the C(5)–C(6) saturated Cyt with the formation of a carbinolamine intermediate and the rate-determining step for deamination at physiological pH is a nucleophilic attack of hydroxide ion on the protonated C(5)–C(6) saturated base [10].

According to several authors (*e.g.* see [3, 10]), the coding properties of such dimers are changed because of deamination of the Cyt base to the corresponding Ura base. If Cyt-containing dimer is not repaired, its Cyt will form the base pair with Ade during the next round of replication. As a result of inefficient repair of damaged DNA by the mismatch repair system the replication errors of nucleic acid lead to G–C→A–T or C→T (U) transition and thus cause a mutation. This idea is supported by the fact that Ade bases are incorporated with high specificity during bypass of site-specific 5′-T[ ]U-3′ or 5′-U[ ]U-3′ dimers [4].

In contrast to the widely discussed UV-mutagenesis mechanism based on the deamination of Cyt-containing photodimers, another mechanism has been proposed

[4,11,12]. According to this mechanism [4,11,12], the saturation of C(5)–C(6) bond of Cyt in Cyt-containing photodimers formed by UV light leads to a shift of the tautomeric equilibrium towards its rare imino form, which becomes the dominant tautomer. In this case, the saturated Cyt has the same base-pairing properties as Thy and Ura, and therefore is able to form a base pair with Ade. In other words, Cyt-containing photodimer directs the incorporation of Ade rather than of the Gua, into the replicating nucleic acid chain. The G–C→A–T (U) or C→T (U) transitions occur as a result of replication errors. In principle, such a change of the nucleotide sequence may lead to the mutation. This is possible when the Cyt imino tautomer is in the *trans* conformation relative to the N(3) atom ring. If the imino tautomer occurs in other geometrically isomeric form (the *cis* conformation), then most likely the Cyt should not be expected to form the base pair with any other base. Because of this, a deletion may occur.

The authors of the experimental work of [4] made conclusion that both the deamination-bypass or tautomer-bypass mechanism could explain the origin of the C→T mutations and the tandem CC→TT mutations. Although the posited UV-mutagenesis mechanism determined by the tautomerism of Cyt-containing dimers explains some of the experimental data (*e.g.* see [4]), there currently does not exist any direct experimental confirmation, nor has a theoretical description of this mechanism been reported to date. In particular, the ability of Pyr[ ]Pyr to adopt the rare tautomeric form(s) remains an unresolved problem, while the question of whether the UV-mutagenesis is a result of the deamination-bypass mechanism or tautomer-bypass mechanism or both mechanisms presents considerable interest.

In our recent works [13–15] it was predicted that the saturation of C(5)–C(6) double bond in Cyt base under formation of ho^{6}hCyt and ^{5,6}hCyt causes the preference of the imino form over the amino form. In addition a quantum mechanical study of pyrimidine photodimers consisting of Cyt bases has been recently performed [16] by the AM1 and PM3 methods. According to these semiempirical calculations it was detected that in the gas phase the preferred form of such cyclobutane dimers includes the rare imino tautomer(s) of Cyt. The imino form of Cyt in the photohydrate or photodimer, which have the same coding properties as Thy and Ura, directs the incorporation of Ade that leads to the G–C→A–T (U) or C→T (U) transitions because of the replication errors. In other words, the data obtained suggest that the nucleic base tautomerism serves as mechanism of UV-induced mutagenesis.

Therefore, to gain an understanding of the molecular mechanisms involved in UV mutagenesis, it is of fundamental interest to estimate the probability of the above-mentioned transformations to occur for the mutagenic effect of photodimers. For this goal it is necessary to study the tendency of the Pyr[ ]Pyr dimer bases occur in their rare form. Since the saturated Cyt is the only pyrimidine base for which the imino tautomer has been found to be more stable than amino one [13–15] the accurate calculations for Cyt-containing pyrimidine dimers are desired. It should be noted that although the semiempirical methods correctly describe the conformation of pyrimidine hydrates [13–15], these methods are not able to predict the correct structure for the pyrimidine photodimers. In particular, the semiempirical methods predict the pla-

nar conformation for the cyclobutyl ring of dimers (*e.g.* see [16,17] and references cited therein) what contradicts the experimental data [1,18]. Such result can be due to the use of zero differential overlap approximation in semiempirical methods. At the same time it was shown (*e.g.* see [19]) that the nonempirical methods turned out to be very effective for the tautomerism of nucleic acids bases and gave excellent results.

It should be noted here that only two articles, [17,20], have been devoted to the study of pyrimidine photodimers by the nonempirical methods. In these articles were presented the results of calculations for the *cis-syn* form of the Ura[ ]Ura and Thy[ ]Thy containing the normal tautomeric forms by the DFT/B3LYP/6-31G*(d) and the HF/6-31G methods, respectively. The results of analogous study for Cyt-containing photodimers representing the greatest interest from the point of view of the presumed role of the cytosine rare tautomer in point mutagenesis are absent.

As part of our ongoing studies of the influence of the saturation of C(5)–C(6) bond of pyrimidine bases on their properties we report in the present paper the results of the theoretical investigation of the energetic properties for all cyclobutane pyrimidine dimers in usual (amino, lactam) and rare (imino, lactim) tautomeric forms. Since the *cis-syn* cyclobutane dimer is predominant form among four possible stereoisomers of Pyr[ ]Pyr and is suggested to be the most mutagenic lesion (*e.g.* see [2,4]), *cis-syn* form was selected to investigate the tautomerism in detail.

METHODS

A quantum mechanical study for Cyt[ ]Cyt, Cyt[ ]Thy and Cyt[ ]Ura photodimers was carried using the *ab initio* LCAO-MO method at the Hartree-Fock (HF) level [21]. In addition, the dimers including one and/or two bases in rare tautomeric forms were also studied, see Scheme 1. The rare imino form of Cyt (Cyt*) was taken in *cis* conformation. These dimers (Cyt*[ ]Cyt, Cyt*[ ]Thy, Cyt*[ ]Ura, Cyt*[ ]Cyt*) are important from the biological point of view due to the possible role of their rare tautomers in point mutagenesis. The other possible location of imino group in regard to N(3) atom of Cyt (the *cis* conformation) have not been considered since it seems not to be important for point mutations.

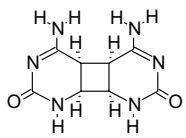
For the molecular orbital expansion we have used the standard 6-31G (d,p) basis set [21], *i.e.* the split valence-shell basis set augmented by a set of six d polarization functions on heavy atoms and three p polarization functions on hydrogens. The tautomeric forms of the pyrimidine photodimer bases under study were optimized without any constraints on their geometry.

In order to estimate the electron correlation contributions to the total electronic energies of the compounds in question the single-point calculations were carried out by means of the Møller-Plesset perturbation theory [22] of the second order (MP2) at the HF-optimized geometries. The core electron correlation effects were not accounted for in these calculations (frozen-core approximation (fc) [21].

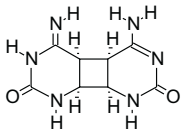
The quantum mechanical calculations were carried out using the GAUSSIAN 98 program [23]. All bond lengths, valence angles and dihedral angles were optimized using the default optimization algorithm (Berny algorithm in redundant internal co-ordinates [24]) and the default convergence criteria set in the Gaussian 98.

Since the Cyt[ ]Cyt photodimer can cause C→T and CC→TT transition mutations the values of total energies for three of most important of Cyt-containing dimers (Cyt[ ]Cyt, Cyt*[ ]Cyt, and Cyt*[ ]Cyt*) have been corrected for the zero-point vibrational energies obtained within the harmonic oscillator approximation at the HF/6-31G(d,p) level of the theory. No scaling factors were applied for the zero-point energy contribution. We do not give the zero-point energy for the other dimers because its contribution to the differences between the energies of dimer containing the rare and normal tautomeric forms of pyrimidine bases is insignificant.

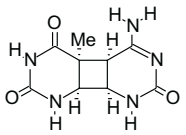
Scheme 1



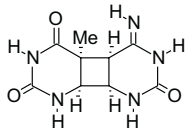
Cyt=Cyt



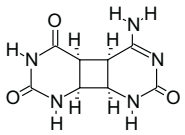
Cyt*=Cyt



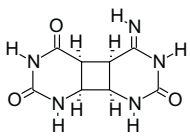
Thy=Cyt



Thy=Cyt*



Ura=Cyt



Ura=Cyt*

It should be noted that the relative stability of tautomeric forms of bases in the dimers found in the gas phase may differ from those observed for molecules as they exist in solution, particularly in a polar solvent. In other words, the tautomeric equilibrium of a molecule depends on the medium in which the tautomerism is studied. A considerable part of the experimental work on the study of physico-chemical and mutagenic properties of the photodimerization reaction was carried out in aqueous solution. In order to study the stability of the compounds in question in water, it is necessary to investigate the influence of an aqueous environment on the tautomeric equilibrium of bases of the pyrimidine photodimers. Similar effects are expected to change the relative order of stabilities of pyrimidine photodimers as a result of interactions with their environment. A correct description of hydration must consider explicitly the interaction between the photodimer and the water molecules. This is a consequence of the fact that the intermolecular interactions are mostly governed by the local specific interactions of polar groups NH_2, NH, C=O, OH with their environment and to a lesser extent by the total dipole moment of the molecule. Thus the water-water and water-dimer interactions on the molecular level should be studied.

Taking into account the above-mentioned considerations, the Monte Carlo simulations of hydration for Cyt[ ]Cyt, Cyt*[ ]Cyt, Cyt*[ ]Cyt* dimers were performed in the canonical (T, V, N) ensemble using Metropolis sampling [25]. In each of the systems studied (dimer + water cluster) 800 water molecules were used. The temperature of the system was 298 K. For the calculation of the intermolecular interaction energies, semiempirical atom-atom potential functions suggested by Poltev *et al*. (see detailed survey of data in [26]) have been used. In these potential functions, the charges assigned to various atoms of the bases and their photohydrates were calculated as we described recently [26]. Monte Carlo calculations were carried out using the gas phase geometry obtained from our an ab initio calculations at Hartree-Fock level, since only small geometric changes were expected upon hydration of rigid molecules of the type considered here.

When periodic boundary conditions are used, the molecules of the solute will be located close to each other even if a relatively large number of the molecules of the solvent is taken into account. In this case,

the simulation will be performed for concentrated solutions. However the solubility of compounds examined is extremely low and, therefore, solutions can be only relatively dilute. Distribution functions can also be distorted due to periodic boundary conditions for a small basic cell. To correctly describe the dilute solution with periodic boundary conditions, it is necessary to use the method of 'the nearest image' and to not take into account the interactions between solute molecules in different independent cells. Instead of the periodic boundary conditions the physical solute-water clusters description (see [27] and references therein) was used in the Monte Carlo simulation. Utilizing the physical cluster theory in the Monte Carlo simulation provides an adequate description for the formation of the nucleation of liquids and "liquid solutions" from vapor. The experimental data which were taken into consideration when the parameters for potential functions were chosen (for details see [26]) were obtained in vacuo by field mass-spectroscopy namely on such conditions.

The system was placed in a sphere with impermeable walls, so that the center of mass of solute coincides with the center of the sphere. The 800 solvent molecules were restricted to remain within a spherical constraining volume centered on the solute that remained fixed in subsequent calculations. A constraining spherical boundary with radius R_c, is centered at the center of mass of the 800 interacting solvent molecules. Initially, the radius of the sphere was set equal to the radius of the sphere for the starting "random" configuration of the cluster of water molecules around each solute R_c = 26.0 Å. These calculations including 4.8×10^6 configurations were performed for the creation of a "water-like" structure inside the sphere. Then according to the definition of the physical cluster [27] for the studied systems the R_c was increased to 31.4 Å and further calculations including 4.8×10^6 configurations were performed. Thus, in order to equilibrate each of the systems, 9.6×10^6 configurations were used, which were then discarded when the average properties were calculated.

In these calculations, the standard deviation statistical error was calculated with the precision of ±0.005. In order to reach the indicated precision for the calculation of the thermodynamic data, the lengths of the generated Markov chain for dimer – water cluster systems were as follows: 5.6×10^7 configurations for Cyt[ ]Cyt, 6.3×10^7 configurations for Cyt*[ ]Cyt, and 5.6×10^7 configurations for Cyt*[ ]Cyt*.

RESULTS AND DISCUSSION

The results of the HF/6-31G**(d,p) optimized structures of homodimers containing the normal tautomeric forms (Cyt[ ]Thy, Cyt[ ]Ura, Cyt[ ]Cyt) and energetically the most preferable heterodimers containing the rare tautomeric form of Cyt (Cyt*[ ]Cyt, Cyt*[ ]Cyt*, Cyt*[ ]Thy, Cyt*[ ]Ura) which cause the C→T and CC→TT transitions are presented in Table 1. The relative internal energies (the relative stabilization energies) for these systems calculated at the HF/6-31G(d,p) and the MP2/6-31G(d,p) levels of theory (ΔE_{HF} and ΔE_{MP2}) are also given in this Table. The relative stabilization energies have been obtained as the difference between the energy (E) for the rare and normal tautomers of these molecules at 0 K. Photodimers containing base molecules in the normal tautomeric forms have been chosen as the reference.

The comparison of the relative stabilization energies for the dimers consisting of Cyt bases (see Table 1) shows that the tautomeric equilibrium of cytosine shifted towards its imino form. *Ab initio* calculations demonstrate that in the gas phase the predominant form of the Cyt in such dimers is the *trans* conformer of its imino form. So according to the prediction of the HF/6-31G(d,p) and MP2/6-31G(d,p)//HF/6-31G(d,p) methods the *trans* conformer of imino tautomer of Cyt[ ]Cyt dimer predominates over its amino one by 5.97 kcal/mol and 5.43 kcal/mol, respectively, *i.e.* the Cyt*[ ]Cyt dimer is more stable than the Cyt[ ]Cyt. The more noticeable difference was detected

for Cyt*[ ]Cyt* dimer. The results of calculations for this dimer led to the total energy difference equal 9.79 and 9.12 kcal/mol by HF/6-31G(d,p) and MP2(fc)//HF/6-31G(d,p) approximations, respectively, in favor of the *trans* imino form of Cyt bases over the amino form of Cyt ones. It is worth noting that only the insignificant part of the preponderance of the *trans* imino tautomer over the amino one is determined by the difference of their zero-point vibrational energies (see Table 1).

Table 1. Energies (a.u.) and relative stabilization energies (kcal/mol) of cytosine-containing pyrimidine photodimers including the normal and rare tautomeric forms calculated at the HF/6-31G(d,p) and the MP2/6-31G(d,p) levels of theory[a].

Photodimer		HF/6-31G(d,p)[b]	MP2/6-31G(d,p)[b]	ΔE_{HF}	ΔE_{MP2}
Cyt[]Cyt	E	–785.232593	–787.605636 (–787.615529)	0.000	0.000 (0.000)
	E(zpe)	–785.013122	–787.386165 (–787.396058)	0.000	0.000 (0.000)
Cyt*[]Cyt	E	–785.242790	–787.614977 (–787.624958)	–6.398	–5.861 (–5.916)
	E(zpe)	–785.022631	–787.394818 (–787.404799)	–5.967	–5.429 (–5.485)
Cyt*[]Cyt*	E	–785.249859	–787.621830 (–787.632120)	–10.834	–10.161 (–10.410)
	E(zpe)	–785.028730	–787.400701 (–787.410991)	–9.793	–9.121 (–9.370)
Cyt[ ]Thy	E	–844.135010	–846.661808	0.000	0.000
Cyt*[ ]Thy	E	–844.141172	–846.668075	–3.866	–3.932
Cyt[ ]Thy*	E	–844.046112	–846.578492	55.780	52.278
Cyt*[ ]Thy*	E	–844.108707	–846.636741	16.504	15.729
Cyt[ ]Ura	E	–805.095928	–807.472123	0.000	0.000
Cyt*[ ]Ura	E	–805.102238	–807.478645	–3.959	–4.092
Cyt[ ]Ura*	E	–805.012390	807.393293	52.417	49.463
Cyt*[ ]Ura*	E	–805.070451	–807.448088	15.986	15.081

[a]Symbol *denotes the rare imino form of cytosine and/or rare lactim form of thymine or uracil; energies E(zpe) are corrected with the zero-point vibrational energy. ΔE_{HF} and ΔE_{MP2} denote the relative internal energy at the 0 K (the difference between the energies of dimer in rare and normal tautomeric forms) calculated by HF and MP2 approximations, respectively. Zero energy corresponds to the normal form of bases of the photodimer. Negative energy values denote a stabilizing effect, while positive energy values denote a destabilizing effect with respect to the normal form.
[b]At HF/6-31G(d,p) optimized geometries.

From Table 1 it is also seen that the similar predominance takes place in the Thy[ ]Cyt and Ura[ ]Cyt heterodimers. As it can be seen these dimers containing the rare form of Cyt are prefered in the gas phase more than dimers containing its normal form. However this predominance is less appreciable than that for Cyt[ ]Cyt dimers (see Table 1).

It should be emphasized that the tautomerism of Cyt[ ]Cyt photodimer is the central for studied problem since this dimer can cause C→T and CC→TT transition mutations. For this reason the MP2 level of theory at the MP2/6-31G(d,p) optimized geometries was also employed for the Cyt[ ]Cyt, Cyt*[ ]Cyt and Cyt*[ ]Cyt* dimers.

The results of our calculations for the compounds in question are given in Table 1 in parentheses.

It is seen that the results obtained for the mentioned dimers by MP2/6-31(d,p)//MP2/6-31(d,p) approximation lead to the same conclusion as that the HF/6-31(d,p) and MP2/6-31(d,p)//HF/6-31(d,p) ones. In other words, the rare tautomer(s) of Cyt in the cytosine photodimer can form a base pair(s) with adenine(s) that can lead to the C→T (U) or CC→TT (UU) transitions because of the replication errors. These data strongly suggest that the tautomerism of Cyt in the Cyt-containing dimers can play an active role in the induction of point mutations.

It should be emphasized that the predicted preference of dimers containing the imino form of Cyt (as compared with the dimers containing its amino form) enables us to make one more important remark. The protonation of the C(5)–C(6) saturated Cyt in the Cyt-containing dimers representing the first step of the hydrolytic deamination reaction [10] most probably proceed not only via the normal tautomeric form of Cyt but also *via* its rare form.

The results reported above refer to isolated molecules (*i.e.* molecules that are not interacting with other molecules of the environment). These results might not, in principle, be correct in aqueous solution, where the majority of the experimental data on the formation of pyrimidine photodimers and their biological role has been obtained. This is because water solvent can stabilize the amino tautomer more than the imino one. Hence the conclusion obtained above may be suspect. Because of this, the study of the amino-imino tautomerism for Cyt-consisting dimers in a water cluster was performed by the Monte-Carlo technique.

The calculated average values of the potential energy for the system (U), the water-water interaction energy (Uww), and the water-base interaction energy (Uwb) for the dimers consisting normal and rare tautomeric forms of the Cyt bases are presented in Table 2. The difference in the internal energy of hydration for these dimers containing imino form(s) of Cyt, relative to those containing amino form of Cyt (ΔE_{aq}), is also given in this Table. Also included here are the values of the standard deviation (statistical error) σ for U and ΔE_{aq}.

As it is shown in Table 2, the Cyt*[ ]Cyt dimer is energetically most preferable in water cluster. The hydration of Cyt*[ ]Cyt dimer is energetically more favorable than hydration of Cyt*[ ]Cyt* one (by 28.8 kcal/mol), *i.e.* the Cyt*[ ]Cyt photodimer is much better hydrated than Cyt*[ ]Cyt*. It follows from Table 2 that a large stabilization of Cyt photodimer including imino form of Cyt upon solvation is completely determined by the U_{ww} term describing the interaction of water molecules with each other. The performed analysis of the radial distribution of the water-water interaction energy shows that water surrounding the Cyt*[ ]Cyt is more structured as compared with that of Cyt[ ]Cyt.

At the same time, the dimer containing both Cyt molecules in the imino form is only slightly stabilized (by 6.4 kcal/mol) with respect to the dimer containing both bases in the amino form. This preference (mentioned above) can also result from the U_{ww} energy term. In the Cyt*[ ]Cyt* dimer case, however, the conclusion about stabil-

ity should be treated with some caution because the magnitute of ΔE_{aq} (see Table 2) exceeds only insignificantly the statistical error appearing when the Monte-Carlo method is used for calculations.

Table 2. Energetic characteristics for the hydration of the Cyt-consisting pyrimidine photodimers including the normal and rare tautomeric forms calculated at the HF/6-31G(d,p) and the MP2/6-31G(d,p) levels of theory[a].

Dimer	$U^b \pm \sigma$	$U_{ww}{}^b \pm \sigma$	$U_{wb}{}^b \pm \sigma$	$\Delta E_{aq}{}^c \pm \sigma$
Cyt[]Cyt	–8.587	–8.386	–0.201	0.0
Cyt*[]Cyt	–8.623	–8.438	–0.185	–28.8 ± 5.6
Cyt*[]Cyt*	–8.595	–8.401	–0.194	–6.4 ± 5.6

[a]Symbol *denotes imino tautomer of cytosine in the dimer.
[b]Kcal/mol water.
[c]Zero energy corresponds to the amino tautomer. Negative energy values denote a stabilizing effect with respect to the amino form. ΔE_{aq} denotes the relative internal energy calculated as the difference between the average potential energies for the rare and normal tautomeric forms of molecule. Owing to the fact that the number of molecules in the water cluster is a constant for the normal and rare tautomers, the changes of the potential energy in the case of tautomeric transformations are equal to the changes in the internal energy. Energies in kcal/mol compound.

It is currently impossible to compare the theoretical results with experiment, because of the absence of the experimental data for simulated systems. However, the quantitative value of this energy preference for Cyt*[ ]Cyt dimer significantly exceeds the Monte Carlo statistical error (see Table 2).

Therefore, the interaction between the water and the Cyt-consisting photodimers does not change the order of the relative stabilities of the most important of the Cyt dimers from that predicted for the gas phase. Moreover, one result of the solvent-induced stabilization for the *trans* conformation is an increase in the relative population of the Cyt dimers including the imino tautomer(s) with respect to the situation in the gas phase.

The data presented above strongly also suggest that, in a water cluster, the Cyt photodimer with one and/or two base in the imino form (its *trans* conformer) is expected to predominate over that with both bases in the amino forms. As a result, the *trans* imino form of Cyt in Cyt-containing dimers can form a pair with Ade that leads to the G–C→A–T or C→T (U) transition because of the replication error. The obtained data suggest of the existence of a mechanism of UV-induced mutagenesis caused by the tautomeric equilibrium shift of Cyt in the Cyt-containing dimers. It follows that these data support the above-mentioned assumption, and serve as evidence that tautomerism of Cyt in the Cyt-containing dimers is a possible mechanism of UV-light mutagenic effect.

It should be emphasized that the study of the mechanism for the mutagenic action of UV light based on tautomerism of the pyrimidine photodimers in DNA is more complicated, and it should include a number of additional factors, which must be taken into consideration. Among them the most significant are as follows. Firstly, an estimation of the influence of the neighboring molecules on a change in the tauto-

meric equilibria needs to be examined. This concerns the study of the tautomeric equilibrium shift due to the influence of the base stacking and hydrogen bonding in DNA. Secondly, the shift conditioned by a change in environment from the isolated base to one surrounded by water molecules should also be studied. The environment of the bases in DNA during several biological processes (replication, translation, transcription) is not exactly the same as that for the bases dissolved in polar solvents. The interior of DNA is supposed to be much less polar than the hydrated DNA exterior. Most likely, the hydration of the DNA bases is local [29]. Anyway, it has been demonstrated that this factor cannot influence the predicted conclusion since the predominant form of Cyt in the Cyt-consisting dimers both in the gas phase and in a water cluster is the imino tautomer. Thirdly, the mismatch formation in DNA does not necessarily lead to a mutation. Presence of modified bases may be responsible for steric or energetic changes in the DNA double strand what in turn may activate the base-excision DNA repair enzymes. These repair enzymes controll the mutation level. In some extreme cases the errors introduced by UV light can be completely removed by repair systems.

CONCLUSIONS

The results obtained by HF and MP2 nonempirical methods indicate that photodimers containing the rare imino form of cytosine are more stable than those containing its normal form. The Monte Carlo simulations of the cytosine-consisting dimers in the water cluster predict that the dimer containing the imino form of cytosine is more stabilized with respect to that containing its amino forms.

Acknowledgments

The nonempirical calculations were carried out in Interdisciplinary Centre for Mathematical and Computational Modelling (ICM) at Warsaw University within the computer grant. The ICM computer centre, Warsaw University, Poland, is acknowledged for grant for the quantum mechanical calculations. One of the authors (V.I.D.) is grateful to Mr. V. Dailidonis for the valuable technical assistance and Mrs. A.V. Golubenko for permanent support and the creation of stimulating environment for this study.

REFERENCES

1. Cadet J. and Vigny P., in Bioorganic Photochemistry, Ed., H. Morrison, NY: John Wiley & Sons, pp. 1–272 (1990).
2. Fisher G.J. and Johns H.E., in Photochemistry and Photobiology of Nucleic Acids, Ed., S.Y. Wang, NY-San Franscisco-London: Acad. Press, v. I, Chemistry, pp. 225–294 (1976).
3. Tessman L., Liu S.K. and Kennedy M.A., *Proc. Natl. Acad. Sci. USA*, **89**, 1159 (1992).
4. Jiang N. and Taylor J.-S, *Biochem*., **32**, 472 (1993).
5. Tu Y., Dammann R. and Pfeifer G.P., *J. Mol. Biol*., **284**, 297 (1998).
6. Hsia H.C., Lebkowski J.S., Leong P-M., Calos M.P. and Miller J.H., *J. Mol. Biol*., **205**, 103 (1989).
7. Keyse S.M., Amaudruz F. and Tyrrell R.M., *Mol. Cell. Biol*., 8, 5425 (1988).
8. Mitchell D.L., *Photochem. Photobiol*., **71**, 162 (2000).

9. Taylor J.-S. and O'Day C.L., *Biochem*., **29**, 1624 (1990).
10. Lemaire D.G.E. and Ruzsicska B.P., *Biochem*., **32**, 2525 (1993).
11. Person S, McCloskey J.A., Snipes W. and Bockrath R.C., *Genetics*, **78**, 1035 (1974).
12. Taylor J.-S., Lu H.-F. and Kotyk J.J., *Photochem. Photobiol*., **51**, 161 (1990).
13. Danilov V.I. and Alderfer J.L., *Biopolymers and Cell (in Russian)*, **15**, 487 (1999).
14. Danilov V.I. and Leś A., *Polish J. Chem*., **74**, 1311 (2000).
15. Danilov V.I., Stewart J.J.P., Leś A. and Alderfer J.L., *Chem. Phys. Lett*., **328**, 75 (2000).
16. Danilov V.I., *Doklady Akademii Nauk Ukrainy (Russian)*, in press.
17. Wiest O. and Largoza A.J., Electrostatic Potential Surfaces for the Cyclobutane Uracil Dimer and Its Radical Ions: Insights into Binding and Catalysis, Internet Journal of Science - Biological Chemistry [on-line], **1**, n.1 [cited 3 March 1997]. Available on the Web: http:/www.netsci-journal.com/97v1/97008/index.htm.
18. Karle I.L., in Photochemistry and Photobiology of Nucleic Acids, Ed., S.Y. Wang, NY-San Franscisco-London: Acad. Press, v. I, Chemistry, pp. 483–519 (1976).
19. Nowak M.J., Lapinski L., Kwiatkowski J.S. and Leszczynski J., in Computational Chemistry: Reviews of Current Trends, Singapore: World Scientific Publishing Co. Pte. Ltd., V.2, pp. 140–216 (1997).
20. Aida M., Kaneko M. and Dupuis M., *Intern. J. Quant. Chem*., **57**, 949 (1996).
21. Hehre W.J., Radom L., Schleyer P.v.R. and Pople J.A., Ab Initio Molecular Orbital Theory, Wiley, NY (1986).
22. Moller-Plesset C., *Phys. Rev.*, **46**, 618 (1934).
23. Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Zakrzewski V.G., Montgomery Jr., J.A., Stratmann R.E., Burant J.C., Dapprich S., Millam J.M., Daniels A.D., Kudin K.N., Strain M.C., Farkas O, Tomasi J, Barone V, Cossi M, Cammi R., Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson G.A, Ayala P.Y., Cui Q., Morokuma K, Malick D.K, Rabuck A.D, Raghavachari K, Foresman J.B., Cioslowski J., Ortiz J.V., Baboul A.G., Stefanov B.B., Liu G., Liashenko A., Piskorz P., Komaromi I., Gomperts R., Martin R.L., Fox D.J., Keith T., Al-Laham M.A., Peng C.Y., Nanayakkara A, Gonzalez C., Challacombe M., Gill P.M.W., Johnson B., Chen W., Wong M.W., Andres J.L., Gonzalez C., Head-Gordon M., Replogle E.S. and Pople J.A. Gaussian 98, Revision A.7. Pittsburgh PA, Gaussian, Inc., 1998.
24. Peng C., Ayala P.Y., Schlegel H.B. and Frisch M.J., *J. Comput. Chem*., **17**, 49 (1996).
25. Metropolis N, Rosenbluth A.W., Rosenbluth M.N., Teller A.N. and Teller E., *J. Chem. Phys*., **21**, 1087 (1953).
26. Alderfer J.L., Danilov V.I., Poltev V.I. and Slyusarchuk O.N., *J. Biomol. Struct. Dyn*., **16**, 1107 (1999).
27. Mruzik M.R., Abraham F.F., Schreiber D.E. and Pound G.M., *J. Chem. Phys*., **64**, 481 (1976).
28. Danilov V.I., Leś A. and Alderfer J.L., *J. Biomol. Struct. Dyn*., in preparation.
29. Schneider B. and Berman H.M., *Biophys. J*., **69**, 2661 (1995).