

Theoretical Study of Light-Induced Crosslinking Reaction Between Pyrimidine DNA Bases and Aromatic Amino Acids

Attila Bende¹*, Alex-Adrian Farcaş^{1,2} and Valer Toşa¹

¹Molecular and Biomolecular Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania, ²Faculty of Physics, Babeş-Bolyai University, Cluj-Napoca, Romania

Low-lying electronic excited states and their relaxation pathways as well as energetics of the crosslinking reaction between uracil as a model system for pyrimidine-type building blocks of DNA and RNA and benzene as a model system for aromatic groups of tyrosine (Tyr) and phenylalanine (Phe) amino acids have been studied in the framework of density functional theory. The equilibrium geometries of the ground and electronic excited states as well as the crossing points between the potential energy surfaces of the uracil-benzene complex were computed. Based on these results, different relaxation pathways of the electronic excited states that lead to either back to the initial geometry configuration or the dimerization between the six-membered rings of the uracil-benzene complex have been identified, and the energetic conditions for their occurrence are discussed. It can be concluded that the DNA-protein crosslinking reaction can be induced by the external electromagnetic field via the dimerization reaction between the six-membered rings of the uracil-benzene pair at the electronic excited-state level of the complex. In the case of the uracil-phenol complex, the configuration of the cyclic adduct (dimerized) conformation is less likely to be formed.

Keywords: DNA-protein crosslinking, cycloaddition reaction, time-dependent density-functional theory, conical intersection, excited-state deactivation pathway, uracil, benzene, phenol

1 INTRODUCTION

The DNA of a living cell is continuously exposed to external influences, which can induce various forms of DNA damage. When this, e.g., appears as a covalent bond of proteins with a DNA strand, through either DNA bases or the sugar-phosphate chain, this phenomenon is called DNA-protein crosslinking (DPC). This is one of the most deleterious forms of DNA damage, which can disturb or even stop the proper cell transcription and replication. To maintain genomic stability, cells must continually repair such damage in a timely manner (Lindahl, 1993; Stingele et al., 2017). DPCs can be induced by exposure to various physical and chemical agents including ionizing radiation, UV light, transition metal ions, environmental contaminants, and common anticancer drugs (Connelly and Leach, 2004; Barker et al., 2005; Ide et al., 2011; Tretyakova et al., 2015).

Electromagnetic radiation such as UV- or visible-light can, through several processes, interact with biological matter and can cause either desired or unwanted biological (Bryant and Frigaard, 2006; Schreier et al., 2007) or chemical changes (Fingerhut et al., 2012; Zhou et al., 2019). DPCs were first recognized as a distinct lesion in UV light-irradiated bacteria by Smith (Smith, 1962) and by

OPEN ACCESS

Edited by:

Carlo Altucci, University of Naples Federico II, Italy

Reviewed by:

Spiridoula Matsika, Temple University, United States Hongmei Su, Beijing Normal University, China

> *Correspondence: Attila Bende

attila.bende@itim-cj.ro

Specialty section:

This article was submitted to Nanobiotechnology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 31 October 2021 Accepted: 16 December 2021 Published: 17 January 2022

Citation:

Bende A, Farcaş A-A and Toşa V (2022) Theoretical Study of Light-Induced Crosslinking Reaction Between Pyrimidine DNA Bases and Aromatic Amino Acids. Front. Bioeng. Biotechnol. 9:806415. doi: 10.3389/fbioe.2021.806415

Alexander and Moroson (Alexander and Moroson, 1962). Biological changes caused by UV radiation have long been considered as harmful as in cases caused by chemical agents, but they were recognized also as the benefit of the phenomenon. Thus, UV crosslinking may be used to selectively label DNAbinding proteins based on their specific interaction with a DNA recognition site (Stützer et al., 2020), and it has been successfully applied to the study of protein–DNA interactions (Welsh and Cantor, 1984; Dimitrov and Moss, 2001). Furthermore, the UVinduced crosslinking and immunoprecipitation (CLIP) technique is used to study dynamical ribonucleoprotein (RNP) assemblies for identifying the sites bound by a specific RNA-binding protein on endogenous RNAs (Spitzer et al., 2014; Hafner et al., 2021).

A simple model system to explore the possibility of light-induced crosslinking reaction between the pyrimidine bases and the side chains of the aromatic amino acids was experimentally investigated by Sun and Fecko (Sun et al., 2006; Fecko et al., 2007). Since the yield of the photoreaction between the pyrimidine bases and the aromatic amino acids is quite low (Shaw et al., 1992; Sun et al., 2006), a short linker of -CH₂- group to mimic the proximity and orientation in DNA-protein complexes was applied between the uracil and benzene fragments. Accordingly, for the final photocyclization of the benzyluracil system, a complex, six-step reaction scheme including electron and proton transfers as well as radical ion pair formation was proposed. Given that these photochemical reactions are very difficult to follow, theoretical modeling can reveal many small details in addition to a global understanding of the processes and thus explain intermediate steps that we cannot observe with our experimental tools (Kleinermanns et al., 2013; González et al., 2016). Accordingly, the model system of benzyluracil was also intensively investigated considering different theoretical frameworks, and several photoreaction mechanisms were proposed (Micciarelli et al., 2013; Micciarelli et al., 2014; Bende and Toşa, 2015; Micciarelli et al., 2017; Valadan et al., 2019; Zhao et al., 2021).

Although the inclusion of the $-CH_2$ - bridge can ensure the proximity of the two fragments, it also determines the outcome of a possible photochemical reaction. In contrast, a simpler photochemical reaction can also take place between two nearby aromatic systems. This phenomenon is known in the literature as thymine dimerization (Boggio-Pasqua et al., 2007; Schreier et al., 2007; Law et al., 2008; Kancheva and Delchev, 2016; Mendieta-Moreno et al., 2016; Zhao et al., 2021). This kind of reaction mechanism does not include any intermediate steps of charge or proton transfer as well as of radical ion pair formation, just a formation of a cyclic adduct (CA) between unsaturated fragments of two molecules.

Our goal is to map using theoretical investigation the possible photochemical processes that may finally occur to light-induced crosslinking reactions between DNA bases and aromatic amino acids.

2 MATERIALS AND METHODS

All equilibrium geometry optimizations and conical intersection (CI) searches were made in the framework of the density functional theory (DFT) considering the ω B97X exchange-

correlation (XC) functional (Chai and Head-Gordon, 2008) combined with the D3-type empirical dispersion correction scheme (Grimme et al., 2010; Lin et al., 2013) and applying the minimally augmented (Zheng et al., 2011) ma-def2-TZVPP triple- ζ basis set of the Karlsruhe group (Weigend and Ahlrichs, 2005) as implemented in the ORCA program suite (Neese, 2012; Neese, 2018). The electronically excited-state calculations were computed using the time-dependent version of the same DFT framework considering the Tamm-Dancoff approximation (TDA) (Hirata and Head-Gordon, 1999). The RIJCOSX approximation (Neese et al., 2009) designed to accelerate the Hartree-Fock and hybrid DFT calculations were considered together with the Def2/J (Weigend, 2006) auxiliary basis set for the Coulomb fitting and def2-TZVPP/C (Hellweg et al., 2007) auxiliary basis set for correlation fitting in the case of TDDFT calculations. The Nudged Elastic Band (NEB) method (Henkelman and Jónsson, 2000; Ásgeirsson et al., 2021) is used to locate the transition state (TS) geometry and to find the minimum energy path (MEP) connecting the minims of different equilibrium geometries on the potential energy surface. The solvent environment of water was taken into account through the conductor-like polarizable continuum (CPCM) model (Barone and Cossi, 1998) and explicitly included two water molecules close to the two carboxyl groups of the uracil molecule. The interaction energies in different molecular dimers were calculated using the pair natural local coupled-cluster method orbital-based (DLPNO-CCSD(T)) (Hansen et al., 2011; Liakos et al., 2011; Riplinger et al., 2013). Based on the localization of the occupied pair natural orbitals, the total energy (reference energy + correlation) can be obtained as a sum of different intra- and inter-fragment contributions. The inter-fragment interaction energy can be written as follows:

$$\Delta E = \delta E_{el.-prep}^{ref.} + E_{elstat.}^{ref.} + E_{exch.}^{ref.} + \delta E_{nonDisp}^{C-CCSD} + E_{Disp}^{C-CCSD} + \delta E_{int.}^{C-(T)}$$
(1)

where $\delta E_{el-prep}^{ref.}$ means the electronic preparation (or intrafragment reference) energy and describes how much energy is necessary to bring the fragments into the electronic structure that is optimal for interaction, $E_{elstat}^{ref.}$, and $E_{exch.}^{ref.}$ are the inter-fragment electrostatic and exchange contributions, $\delta E_{nonDisp}^{C-CCSD}$ and E_{Disp}^{C-CCSD} are the non-dispersive and dispersive parts of the correlation energy at the CCSD level, and $\delta E_{int.}^{C-(T)}$ is the triples correction term to the inter-fragment interaction energy. The electronically excited states were also computed considering the SCS-PBE-QIDH XC functional (Casanova-Páez and Goerigk, 2021) built as the spin-component scaled version of Adamo's PBE-based double hybrid (Brémond et al., 2014) optimized for excited states, all of them implemented in the same orca package. Although the TDDFT method can accurately describe the equilibrium geometries of excited states and their energies, it cannot accurately determine the CI points due to the problem of the wrong dimensionality of the CI between ground and excited states at the TDDFT level (Levine et al., 2006; Huix-Rotllant et al., 2016; Huix-Rotllant et al., 2020). Accordingly, in addition to the standard linear-response TDDFT, CI points were also



determined by the spin-flip TDDFT (or SF-TDDFT) method (Rinkevicius et al., 2010; Mei and Yang, 2019; Casanova and Krylov, 2020; Matsika, 2021) implemented in the same Orca package. The molecular geometries were built, analyzed, and further manipulated using the Gabedit (Allouche, 2011) and Avogadro (Hanwell et al., 2012), while the molecular graphics were created using the GaussView (Dennington et al., 2009) software.

3 RESULTS AND DISCUSSIONS

3.1 Benzene–Uracil Model

The benzene-uracil (B-U) molecular complex can be considered as a model system for describing the light-induced crosslinking reaction between the thymine DNA base (or uracil RNA base) and the phenylalanine-type aromatic side chain of a protein. Accordingly, in the first step, the stacking π – π configuration was built between the aromatic rings of the benzene and uracil. In the case of the benzene ring, the -CH2-CH3 fragment was also added to mimic the amino acid side chain. Performing geometry optimization without any constrain, using the wB97X-D3/madef2-TZVPP level of theory, the parallel-shifted stacking configuration was obtained. The molecular graphics is presented in Figure 1, where in panel A is shown the side view, while in panel B, the top view of the molecular dimer is given. It is important to note that the stacking configuration of the two aromatic rings is not perfect plane-parallel; the shortest interplanar distance is 3.435 Å, while the largest one is 3.640 Å. The intermolecular interaction energy between the benzene and uracil units is -10.75 kcal/mol obtained considering the same level of theory used for the geometry optimization. The energy value obtained at the DFT level matches very well with the energy of -11.17 kcal/mol computed considering the DLPNO-CCSD(T) theory. To understand in more detail the nature of the intermolecular interaction, the energy decomposition scheme was applied in the case of DLPNO-CCSD(T) theory. Accordingly, the strength of the interaction without any electron correlation effects is positive ($\Delta E^{HF} = 2.39 \text{ kcal/mol}$), the stacking configuration is stabilized by electron correlation, where the interaction at CCSD (coupled-cluster including singles and doubles cluster operators) is -10.43 kcal/mol, and

-11.17 kcal/mol if the perturbative triplet excitation to the cluster expansion is taken into account. The results of the energy decomposition analysis show that the electron correlation effects are mainly covered by the dispersion effects $(E_{Disp}^{C-CCSD} = -12.28 \text{ kcal/mol})$ characteristic for the π - π stacking interaction, which is slightly enhanced by the non-dispersion $(\delta E_{nonDisp}^{C-CCSD} = -0.55 \text{ kcal/mol})$ and the triples correction $(\delta E_{int.}^{C-CCT}) = -0.73 \text{ kcal/mol})$ effects. As can be observed, the dispersion electron correlation effects are decisive in the formation of the stacking configuration, and therefore, theoretical methods that do not take these effects into account cannot correctly describe either the geometry or its electron configuration. Accordingly, to describe electron excitations and their effects on molecular geometry, it is also necessary to use a method that takes these dispersion effects into account.

To better understand the structural transition between the stacking-type aggregated complex of the B-U complex and the bounded B-U complex one, not only the "reactants" but also the "product" structure needs to characterize, too. Accordingly, the equilibrium geometry of the CA structure was computed using the same wB97X-D3/ma-def2-TZVPP level of theory followed by the NEB-type calculation for finding the TS geometry between the stacked and CA geometries. The graphics of the CA and TS geometries are compiled in Figure 2. In the CA configuration case, two saturated bond formations can be observed, through which the aromatic nature of the benzene ring is broken. The C-C bond lengths between the two molecular rings are equivalent, each of them having 1.565 Å of bond distance, while the lengths of corresponding double bonds are modified from 1.390 to 1.552 Å in the case of the benzene and from 1.343 to 1.559 Å for the uracil's hexagonal ring. In the case of the TS geometry, only a single C-C bond of 1.567 Å length is formed between the two rings, but the aromatic of the benzene and the double-bond character of the uracil rings remains broken as it was found for the CA geometry. The conformation energy ($\Delta E = E^{stack} - E^{CA}$) between the stacking and CA geometries is -29.23 kcal/mol, while the "left" barrier ($\Delta E = E^{stack} - E^{TS}$) between the stacking and TS geometries is -70.33 kcal/mol as well as the "right" barrier $[(\Delta E = E^{CA} - E^{TS})]$ between the CA and TS geometries is -41.10 kcal/mol. Both "left"- and "right"-barrier energy values are given by the TS geometry, indicating that thoroughfare between the two final structures is less likely





since the barriers are high enough to keep stable the two equilibrium geometries, and it is even less likely that stacking geometry will flip over in CA geometry. Accordingly, it can be concluded that in the ground state, the stacking B–U complex is not able to change into the CA-type dimerized configuration.

To check whether this geometry transition can occur at the electronic excited-state level, TDDFT analysis was initiated. Accordingly, the first three, low-lying electronic excited states for the B-U stacking complex computed at TD-ωB97X-D3/madef2-TZVPP level of theory are $S_1 = 233$ nm, $S_2 = 226$ nm, and S_3 = 218 nm. Since, in the case of ω B97X XC functional, the dispersion effects are not explicitly taken into account in the correlation part of the functional that is only an empirical postcorrection to the energy calculation, the double-hybrid SCS-PBE-QIDH XC functional was also considered. The same first three low-lying electronic excited states computed at this time using the TD-SCS-PBE-QIDH/ma-def2-TZVPP level of theory are $S_1 =$ 241 nm, $S_2 = 226$ nm, and $S_3 = 214$ nm. The theoretical UV spectra of the B-U complex obtained considering the ωB97X-D3 (left panel) and SCS-PBE-QIDH (right panel) XC functionals are shown in Figure 3, while the Natural Difference Orbitals (NDOs) computed for the first three low-lying excited states of the B-U complex using the TD-wB97X-D3/ma-def2-TZVPP level of theory are presented in Figure 4. The results presented in Figures 3, 4 show that the S_1 electronic excited state is mainly

localized on the uracil, the S_2 state on the benzene, and the S_3 state again on the uracil unit of the B–U complex. It is also important to note that among the three electronic excited states, only the S_1 state presents significant oscillator strength and accordingly efficient absorption of electromagnetic radiation.

To get a much more comprehensive picture of the absorption of electromagnetic radiation, it is worth comparing the B-U binary system with the isolated uracil and benzene cases themselves. The low-lying excited states of the uracil were investigated in detail by Epifanovsky et al. (2008), showing that in the gas phase the $S_0 \rightarrow S_1$ excitation is a dark, electronically forbidden $n \rightarrow \pi^*$ transition with $\omega = 237$ nm (f = 0.0) frequency obtained at EOM-CCSD/aug-ANO-DZ (equation-of-motion coupled-cluster with single and double excitations) level of theory, while the second $S_0 \rightarrow S_2$ excitation with $\pi \to \pi^{\star}$ character has a very strong absorption intensity with $\omega = 222 \text{ nm} (f = 0.2110)$ frequency. These values are in a very good agreement with the present results obtained at TD- ω B97X-D3/ma-def2-TZVPP level of theory, showing S₁ = 237 nm (f = 0.00003) and S₂ = 223 nm (f = 0.2682) excitation energy values. On the other hand, the inclusion of the solvent (water) affects the excitation energies of the two excited states, where the order of the two excited states, which were already close to each other in the gas phase, is simply reversed, and the energy values are blue-shifted: $S_1 = 228 \text{ nm} (f = 0.3485) \text{ and } S_2 = 219 \text{ nm}$







(f = 0.0001). Regarding the solvent effect, a similar phenomenon can be observed for the binary system B–U. In the gas phase, the two electronic transitions are S₁ = 233 nm (f = 0.00005) and S₂ = 226 nm (f = 0.0616) obtained at TD- ω B97X-D3/ma-def2-TZVPP level of theory, while in water using CPCM solvent model, they are S₁ = 230 nm (f = 0.2295) and S₂ = 225 nm (f = 0.0064), with the difference that it is not the two excited states localized on uracil that is being reversed but one excited state of the uracil with one of the benzene.

In summary, the analysis of the vertical excitation energies indicates the fact that inside the B-U binary complex, the lowest energy of the electromagnetic radiation is absorbed by the uracil units. The next step would be the analysis of the first electronic excited-state relaxation pathway. Since no real adiabatic dynamics is performed, using a simple energy optimization technique, one can follow only the energetic behavior of the relaxation path. Accordingly, starting the geometry optimization from the ground state geometry configuration and using the same TD-wB97X-D3/ma-def2-TZVPP level of theory, there are stable intermediate geometries (not an energetically conformation) where the two hexagonal rings get closer to each other and the planar shift of the rings vanishes, as well as the NDO already overlaps the two hexagonal rings. For molecular graphics of one of these geometries, see Figure 5A, while the NDO of its $S_0 \rightarrow S_1$ transition is presented in **Figure 5B**.

This particular intermediate geometry was chosen only to illustrate how the excited electron charge distribution changes during relaxation. Accordingly, the two rings start to close each other, and the excited electron that was initially localized on the uracil spreads also over the benzene ring, inducing a strong charge transfer. For the chosen intermediate geometry, this charge transfer is 0.24 e, moving from B to U (based on the Löwdin population analysis). Finally, the optimization of the S₁ excited-state energy falls close to an $(S_0 \otimes S_1)_a$ CI point. In the next step, considering as starting geometry the configuration of the last geometry step during the previous S₁ excited-state optimization, a searching procedure of the CI was performed. The geometry of the successfully localized CI point is shown in **Figure 6A**.

During the first excited-state geometry optimization, the B–U binary complex arrives at a crossing point between the S_1 excited and ground states potential energy surfaces, where it loses its excited-state character and turns to the ground-state electron configuration. Performing NEB studies on the nature of the potential energy surface on the one hand between the stacking equilibrium and the $S_0 \otimes S_1$ CI geometries and on the other hand between the CA equilibrium and the $S_0 \otimes S_1$ CI geometries, no TS geometry was found in either case. This means that after the changing of the S_1 electronic excited state to the S_0 ground state, the system almost randomly can choose its relaxation path either to the stacking or to the CA geometry. In this way, it can be stated





that the excitation considering the first excited state of the stacking configuration of the B-U binary complex finally leads either to back the stacking configuration via the radiationless deexcitation or to the CA configuration via the dimerization reaction. Due to the high enough energy barrier, the CA configuration remains stable, which means that the external electromagnetic radiation to some extent can induce DPC via the CA dimerization. It is well-known that due to the wrong dimensionality of the branching space, the standard TDDFT breaks down near an $S_0 \otimes S_1$ and fails to correctly compute the CI geometries. On the other hand, it was already proven that dynamic electron correlation counted by dispersion effects is essential to incorrectly describe the supramolecular complex. Therefore, simply applying the multireference HF methods without any perturbational or configuration interaction type correction will also lead to an incorrect description. In this respect, the SF approach of the TDDFT provides a more balanced description of the topology of the branching space. Accordingly, the $(S_0 \otimes S_1)_a$ geometry was reoptimized considering the SF-TDDFT method. The structural deviation between the standard and the SF versions of the TDDFT is shown in Figure 7A. The major difference can be observed in the orientation of the uracil fragment compared with the benzene

unit. More precisely, the new bond formed between the uracil and benzene units changes from the 1.659 Å obtained for the initial case to 1.591 Å for the SF case, as well as rotates the uracil fragment with an additional 20° relative to the benzene ring. This additional rotation means that the C–C bond distance, in which the further decrease is necessary to obtain the CA configuration, will be larger with more than 0.3 Å, and thus, the likelihood of CA configuration also decreases. Furthermore, from the energetic point of view, SP-TDDFT also reduces the height of the $\Delta(E^{stack} - E^{CI})$ energy barrier by ≈20 kcal/mol. Despite the differences, the standard TDDFT can also estimate a near-good S₀ ⊗S₁ geometry, but of course, a more accurate method, like SF-TDDFT, is needed to describe them accurately.

Another crossing point between the S_1 excited and ground state on the potential energy hyper-surface was also found, which, in turn, is unlikely to lead to the formation of a CA configuration. The molecular graphics of this CI crossing point geometry is presented in **Figure 6B**, where it can be seen that one of the carboxyl groups from the uracil units enlarges its C=O bond, turning to a single bond, and the oxygen makes a new covalent bond with one of the carbon atoms from the benzene ring. The length of the C^{*U*}-O bond obtained with the standard TDDFT method is 1.417 Å, while that of the C^{*B*}-O bond becomes 1.458 Å.





This CI geometry is a very interesting configuration because if one starts a geometric optimization from this position, the system turns to the stacking conformation. This conformational state is located at an energy value of 166.2 kcal/mol higher than the energy of the stacking geometry. The interesting thing about both CI cases is that not only the uracil or benzene ring is formed, but in both cases, the joint deformation of the two rings can be observed. As seen earlier, in addition to the standard TDDFT, another more advanced method is needed to accurately describe the CI geometry. Accordingly, the $(S_0 \otimes S_1)_b$ geometry was again reoptimized, considering the SF-TDDFT method. The structural deviation between the standard and the SF versions of the TDDFT is shown in Figure 7B. The results show that the length of the C^U -O bond decreases to 1.333 Å, and that of the C^{B} –O bond increases to 1.524 Å compared with the initial geometry parameters, while the corresponding energy barrier decreases with \approx 57 kcal/mol.

Another important aspect is what other kinds of photoproduct geometries can be formed after the irradiation and how energetically they seem to be real. Accordingly, new binary complexes considering different direct bonding possibilities were built, and their geometries were optimized. In the first complex, the uracil and benzene rings are bounded through the C–C single covalent bond (**Figure 8A**), while in the second case, the O atom of the uracil makes a covalent bond with a C atom from the benzene ring (Figure 8B). It is important to note that in both cases, the formation of photoproducts takes place through several reaction steps. While in the first case we are talking about a two-step process where, in addition to the formation of the covalent bond, dehydrogenation also takes place, in the second case, an intermediate step of a bond conjugation also happens. If one compares these geometries energetically to the stacking configuration, it can be found that the conformational energy difference for case *a*) is +19.17 kcal/mol ($E^{diff} = +29.23$ kcal/mol) while for case b) is +45.98 kcal/mol. The geometry found in the first case seems to be more stable even compared with the CA configuration, but it is a question of what successive energy barriers need to be overcome for the reaction to take place. One possible turning point is the formation of the $(S_0 \otimes S_1)_b$ CI point (Figure 6B) since in this case the aromatic nature of the benzene ring is already broken, and it may be possible for dehydrogenation to take place.

3.2 Phenol–Uracil Model

Considering the phenol–uracil (P–U) model system, one can model the light-induced crosslinking reaction for the phenylalanine–uracil protein–DNA cases. The geometry optimization of the P–U binary complex results in a quasi-stacking parallel configuration where the tilted position is induced by the hydrogen-bond network made by the three water, the O atom of the uracil, and the OH fragment of the









FIGURE 12 | The potential energy profiles including equilibrium geometry, vertical excitation, transition state, and conical intersection point energies for the benzene–uracil (B–U) and phenol–uracil (P–U) complexes. Energy values of different conformational barriers are given in kcal/mol, while those in parentheses are in eV. Red path color means the electronic excited-state relaxation, while the blue one represents the ground electronic state relaxation.

phenol units (**Figure 9A**). In this case, the water molecules were also depicted in the molecular graphics. The shortest inter-planar distance (O...O) is 3.490 Å, while the largest one is 3.870 Å. The intermolecular interaction energy obtained at the DFT level of theory is -7.87 kcal/mol, while that found considering the DLPNO-CCSD(T) theory is -8.47 kcal/mol. The stacking configuration is exclusively stabilized by the dispersion-type electron correlation effects, in which their contributions to the final intermolecular binding energy are -8.89 kcal/mol.

The first low-lying electronic excited states were computed using the ω B97X-D3 and SCS-PBE-QIDH XC functionals. The excitation energies are as follows: S₁ = 238 nm, S₂ = 231 nm, and S₃ = 223 nm in the case of ω B97X-D3 and S₁ = 254 nm, S₂ = 240 nm, and S₃ = 226 nm for the SCS-PBE-QIDH double-hybrid XC functional. The most significant difference from the previously studied system is that in the present case the first excited state is localized on the phenol unit and not on the uracil component as one was found for the B–U case. The second

electronic excited state is almost entirely uracil in nature, while the third one is a $P \rightarrow U$ charge transfer state. For their NDOs, see **Figure 10**.

After the calculation of the vertical excitation energies, the equilibrium geometry of the excited state S_1 was also determined. The molecular graphics of the S_1 equilibrium geometry is shown in **Figure 11**. The geometric changes compared with the ground state geometry are manifested in the fact that the two hexagon rings get closer (the closest interatomic distance is 2.927 Å) and rotate in the plane at 60° relative to each other. It is characteristic of both the P–U and B–U systems that the distance between the stacking planes decreases during the S_1 relaxation. From this point of view, it may even seem irrelevant on which fragment the vertical excitation is located.

Similar to the B-U system, the CA-type equilibrium geometry of the P-U binary complex was also computed, and its geometry is shown in Figure 9B. Unlike the B-U system, in the P-U binary complex case, the C-C vertical bond lengths between the two molecular rings are not equivalent (1.562 and 1.587 Å); the carbon atom of the benzene ring that binds also the OH fragment makes a bit larger C-C vertical bond than the other C-C bond. The conformational energy is 34.53 kcal/mol higher compared with the stacking system. With the use of the NEB TS geometry searching method, the TS geometry between the stacking and CA configurations was also computed. The "left" barrier ($\Delta E = E^{stack} - E^{TS}$) between the stacking and TS geometries is -44.95 kcal/mol as well as the "right" barrier (ΔE $= E^{CA} - E^{TS}$) between the CA and TS geometries is only -10.43 kcal/mol. This latter value means that the CA geometry is no longer as stable as it was found for the B-U system. In this way, the dimerized conformation (CA) and implicitly the crosslinked P-U configuration are less probable to be formed, and thus during a possible CI-driven relaxation, the probability of CA geometry appearing is lower. Of course, all these findings are true when the carbon atom to which the phenyl group is bound takes part in the CA formation. When this is not the case, namely, the CA is formed between two benzene-like carbon atoms, then, of course, the configuration already analyzed for the B-U system will be valid with a good approximation.

Analyzing the potential energy profiles both for the B–U and for the P–U cases shown in **Figure 12**, it can be said that relaxation after vertical excitation of either stacking or CA structures very easily leads to the $(S_0 \otimes S_1)_a$ CI geometry and from there returns to the ground state electron structure of one of the structures without spending much time in their excited state.

4 CONCLUSION

In the present work, considering the B–U and phenol–uracil model systems, the light-induced DPC reaction was investigated based on the ω B97X-D3 and SCS-PBE-QIDH XC DFT functionals as well as DLPNO-CCSD (T) coupled-cluster theory. The main focus was oriented on the possible occurrence of CA reactions similar to thymine dimerization in DNA (Mendieta-Moreno et al., 2016). The result obtained for the first excited-state relaxation pathway of B–U shows that the S₁ state presents

a spontaneous relaxation till it reaches the supramoleculartype $(S_0 \otimes S_1)_a$ CI point, defined by the deformation of both uracil and benzene moieties, from which the B-U binary system slides almost randomly into either stacking or CAtype geometry. From this, it can be concluded that during the photo-reaction induced by the external electric field, both biologically positive internal conversion, serving as a protection mechanism against the UV radiation, and biologically negative dimerization can occur as possible scenarios. In the case of the phenol-uracil binary system, where the carbon atom that holds the OH group is involved in the CA-type binding, the occurrence of the CA product is less possible, whereas the energy barrier between the TS and the CA-type geometries is much lower than that found for the B-U system. This can certainly be explained by the presence of the OH group in the phenol unit. At the same time, we also found another supramolecular-type CI configuration $(S_0 \otimes S_1)_b$, which in our view could be the starting point of the multistep reaction during which crosslinking configurations known in the literature (Spitzer et al., 2014; Hafner et al., 2021) are formed.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization: AB. Methodology: AB. Validation: AB and VT. Investigation: AB and A-AF. Writing—original draft preparation: AB and VT. Writing—review and editing: AB and VT. Visualization: AB and A-AF. Supervision: AB and VT. Project administration: AB. Funding acquisition: AB. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the UEFISCDI public institution under the Romanian Ministry of Education, Project Code: PN-III-P4-ID-PCE-2020-0770. The article processing charge (APC) was funded by the Ministry of Research, Innovation and Digitization through Programme 1—Development of the National Research and Development System, Subprogram 1.2—Institutional Performance—Funding Projects for Excellence in RDI, Contract No. 37PFE/30.12.2021.

ACKNOWLEDGMENTS

The authors thank INCDTIM, Cluj-Napoca Data Center, for providing computer facilities.

REFERENCES

- Alexander, P., and Moroson, H. (1962). Cross-linking of Deoxyribonucleic Acid to Protein Following Ultra-violet Irradiation of Different Cells. *Nature* 194, 882–883. doi:10.1038/194882a0
- Allouche, A.-R. (2011). Gabedit-A Graphical User Interface for Computational Chemistry Softwares. J. Comput. Chem. 32, 174–182. doi:10.1002/jcc.21600
- Ásgeirsson, V., Birgisson, B. O., Bjornsson, R., Becker, U., Neese, F., Riplinger, C., et al. (2021). Nudged Elastic Band Method for Molecular Reactions Using Energy-Weighted Springs Combined with Eigenvector Following. J. Chem. Theor. Comput. 17, 4929–4945. doi:10.1021/acs.jctc.1c00462
- Barker, S., Weinfeld, M., and Murray, D. (2005). DNA-protein Crosslinks: Their Induction, Repair, and Biological Consequences. *Mutat. Research/Reviews Mutat. Res.* 589, 111–135. doi:10.1016/j.mrrev.2004.11.003
- Barone, V., and Cossi, M. (1998). Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. J. Phys. Chem. A. 102, 1995–2001. doi:10.1021/jp9716997
- Bende, A., and Toşa, V. (2015). Modeling Laser Induced Molecule Excitation Using Real-Time Time-dependent Density Functional Theory: Application to 5- and 6-benzyluracil. *Phys. Chem. Chem. Phys.* 17, 5861–5871. doi:10.1039/ c4cp03869j
- Boggio-Pasqua, M., Groenhof, G., Schäfer, L. V., Grubmüller, H., and Robb, M. A. (2007). Ultrafast Deactivation Channel for Thymine Dimerization. J. Am. Chem. Soc. 129, 10996–10997. doi:10.1021/ja073628j
- Brémond, É., Sancho-García, J. C., Pérez-Jiménez, Á. J., and Adamo, C. (2014). Communication: Double-Hybrid Functionals from Adiabatic-Connection: The QIDH Model. J. Chem. Phys. 141, 031101. doi:10.1063/1.4890314
- Bryant, D. A., and Frigaard, N.-U. (2006). Prokaryotic Photosynthesis and Phototrophy Illuminated. *Trends Microbiol.* 14, 488–496. doi:10.1016/ j.tim.2006.09.001
- Casanova, D., and Krylov, A. I. (2020). Spin-flip Methods in Quantum Chemistry. Phys. Chem. Chem. Phys. 22, 4326–4342. doi:10.1039/C9CP06507E
- Casanova-Páez, M., and Goerigk, L. (2021). Time-Dependent Long-Range-Corrected Double-Hybrid Density Functionals with Spin-Component and Spin-Opposite Scaling: A Comprehensive Analysis of Singlet-Singlet and Singlet-Triplet Excitation Energies. J. Chem. Theor. Comput. 17, 5165–5186. doi:10.1021/acs.jctc.1c00535
- Chai, J.-D., and Head-Gordon, M. (2008). Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. J. Chem. Phys. 128, 084106. doi:10.1063/ 1.2834918
- Connelly, J. C., and Leach, D. R. F. (2004). Repair of DNA Covalently Linked to Protein. *Mol. Cel* 13, 307–316. doi:10.1016/S1097-2765(04)00056-5
- Dennington, R., Keith, T., and Millam, J. (2009). *Gaussview 5.0.9.* Shawnee Mission, KS: Semichem Inc.
- Dimitrov, S. I., and Moss, T. (2001). UV Laser-Induced Protein-DNA Crosslinking. In DNA-Protein Interactions. *Methods Mol. Biol.* 148, 395–402. doi:10.1385/1-59259-208-2:395
- Epifanovsky, E., Kowalski, K., Fan, P.-D., Valiev, M., Matsika, S., and Krylov, A. I. (2008). On the Electronically Excited States of Uracil. J. Phys. Chem. A. 112, 9983–9992. doi:10.1021/jp803758q
- Fecko, C. J., Munson, K. M., Saunders, A., Sun, G., Begley, T. P., Lis, J. T., et al. (2007). Comparison of Femtosecond Laser and Continuous Wave UV Sources for Protein-Nucleic Acid Crosslinking. *Photochem. Photobiol.* 83, 1394–1404. doi:10.1111/j.1751-1097.2007.00179.x
- Fingerhut, B. P., Herzog, T. T., Ryseck, G., Haiser, K., Graupner, F. F., Heil, K., et al. (2012). Dynamics of Ultraviolet-Induced DNA Lesions: Dewar Formation Guided by Pre-tension Induced by the Backbone. *New J. Phys.* 14, 065006. doi:10.1088/1367-2630/14/6/065006
- González, J., Baños, I., León, I., Contreras-García, J., Cocinero, E. J., Lesarri, A., et al. (2016). Unravelling Protein-DNA Interactions at Molecular Level: A DFT and NCI Study. J. Chem. Theor. Comput. 12, 523–534. doi:10.1021/ acs.jctc.5b00330
- Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. (2010). A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. J. Chem. Phys. 132, 154104. doi:10.1063/1.3382344

- Hafner, M., Katsantoni, M., Köster, T., Marks, J., Mukherjee, J., Staiger, D., et al. (2021). Clip and Complementary Methods. *Nat. Rev. Methods Primers* 1, 20. doi:10.1038/s43586-021-00018-1
- Hansen, A., Liakos, D. G., and Neese, F. (2011). Efficient and Accurate Local Single Reference Correlation Methods for High-Spin Open-Shell Molecules Using Pair Natural Orbitals. J. Chem. Phys. 135, 214102. doi:10.1063/1.3663855
- Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012). Avogadro: an Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform* 4, 17. doi:10.1186/1758-2946-4-17
- Hellweg, A., Hättig, C., Höfener, S., and Klopper, W. (2007). Optimized Accurate Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations for the Atoms Rb to Rn. *Theor. Chem. Acc.* 117, 587–597. doi:10.1007/s00214-007-0250-5
- Henkelman, G., and Jónsson, H. (2000). Improved tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. J. Chem. Phys. 113, 9978–9985. doi:10.1063/1.1323224
- Hirata, S., and Head-Gordon, M. (1999). Time-dependent Density Functional Theory within the Tamm-Dancoff Approximation. *Chem. Phys. Lett.* 314, 291–299. doi:10.1016/S0009-2614(99)01149-5
- Huix-Rotllant, M., Ferré, N., and Barbatti, M. (2020). Time-Dependent Density Functional Theory. Quant. Chem. Dyn. Excited States 2, 13–46. doi:10.1016/ S0065-3276(08)60600-0
- Huix-Rotllant, M., Nikiforov, A., Thiel, W., and Filatov, M. (2015). Description of Conical Intersections with Density Functional Methods. *Top. Curr. Chem.* 368, 445–476. doi:10.1007/128_2015_631
- Ide, H., Shoulkamy, M. I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A. M. H. (2011). Repair and Biochemical Effects of DNA-Protein Crosslinks. *Mutat. Research/Fundamental Mol. Mech. Mutagenesis* 711, 113–122. doi:10.1016/ j.mrfmmm.2010.12.007
- Kancheva, P. B., and Delchev, V. B. (2016). Investigation of the Mechanisms of Photo-Induced Formation of Cyclobutane Dimers of Cytosine and 2,4diaminopyrimidine. J. Mol. Model. 22, 230. doi:10.1007/s00894-016-3087-9
- Kleinermanns, K., Nachtigallová, D., and de Vries, M. S. (2013). Excited State Dynamics of DNA Bases. *Int. Rev. Phys. Chem.* 32, 308–342. doi:10.1080/ 0144235X.2012.760884
- Law, Y. K., Azadi, J., Crespo-Hernández, C. E., Olmon, E., and Kohler, B. (2008). Predicting Thymine Dimerization Yields from Molecular Dynamics Simulations. *Biophysical J.* 94, 3590–3600. doi:10.1529/biophysj.107.118612
- Levine, B. G., Ko, C., Quenneville, J., and MartÍnez, T. J. (2006). Conical Intersections and Double Excitations in Time-dependent Density Functional Theory. *Mol. Phys.* 104, 1039–1051. doi:10.1080/00268970500417762
- Liakos, D. G., Hansen, A., and Neese, F. (2011). Weak Molecular Interactions Studied with Parallel Implementations of the Local Pair Natural Orbital Coupled Pair and Coupled Cluster Methods. J. Chem. Theor. Comput. 7, 76–87. doi:10.1021/ct100445s
- Lin, Y.-S., Li, G.-D., Mao, S.-P., and Chai, J.-D. (2013). Long-range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. J. Chem. Theor. Comput. 9, 263–272. doi:10.1021/ct300715s
- Lindahl, T. (1993). Instability and Decay of the Primary Structure of DNA. *Nature* 362, 709–715. doi:10.1038/362709a0
- Matsika, S. (2021). Electronic Structure Methods for the Description of Nonadiabatic Effects and Conical Intersections. *Chem. Rev.* 121, 9407–9449. doi:10.1021/acs.chemrev.1c00074
- Mei, Y., and Yang, W. (2019). Excited-State Potential Energy Surfaces, Conical Intersections, and Analytical Gradients from Ground-State Density Functional Theory. J. Phys. Chem. Lett. 10, 2538–2545. doi:10.1021/ acs.jpclett.9b00712
- Mendieta-Moreno, J. I., Trabada, D. G., Mendieta, J., Lewis, J. P., Gómez-Puertas, P., and Ortega, J. (2016). Quantum Mechanics/molecular Mechanics Free Energy Maps and Nonadiabatic Simulations for a Photochemical Reaction in DNA: Cyclobutane Thymine Dimer. J. Phys. Chem. Lett. 7, 4391–4397. doi:10.1021/acs.jpclett.6b02168
- Micciarelli, M., Altucci, C., Della Ventura, B., Velotta, R., Toşa, V., Pérez-Góonzalez, A. B., et al. (2013). Low-lying Excited-States of 5-benzyluracil. *Phys. Chem. Chem. Phys.* 15, 7161–7173. doi:10.1039/C3CP50343G
- Micciarelli, M., Curchod, B. F. E., Bonella, S., Altucci, C., Valadan, M., Rothlisberger, U., et al. (2017). Characterization of the Photochemical

Properties of 5-benzyluracil via Time-dependent Density Functional Theory. J. Phys. Chem. A. 121, 3909–3917. doi:10.1021/acs.jpca.6b12799

- Micciarelli, M., Valadan, M., Della Ventura, B., Di Fabio, G., De Napoli, L., Bonella, S., et al. (2014). Photophysics and Photochemistry of a DNA-Protein Cross-Linking Model: A Synergistic Approach Combining Experiments and Theory. J. Phys. Chem. B 118, 4983–4992. doi:10.1021/jp4115018
- Neese, F. (2018). Software Update: the Orca Program System, Version 4.0. Wires Comput. Mol. Sci. 8, e1327. doi:10.1002/wcms.1327
- Neese, F. (2012). The Orca Program System. Wires Comput. Mol. Sci. 2, 73–78. doi:10.1002/wcms.81
- Neese, F., Wennmohs, F., Hansen, A., and Becker, U. (2009). Efficient, Approximate and Parallel Hartree-Fock and Hybrid DFT Calculations. A 'chain-Of-Spheres' Algorithm for the Hartree-Fock Exchange. *Chem. Phys.* 356, 98–109. doi:10.1016/j.chemphys.2008.10.036
- Rinkevicius, Z., Vahtras, O., and Ågren, H. (2010). Spin-flip Time Dependent Density Functional Theory Applied to Excited States with Single, Double, or Mixed Electron Excitation Character. J. Chem. Phys. 133, 114104. doi:10.1063/ 1.3479401
- Riplinger, C., Sandhoefer, B., Hansen, A., and Neese, F. (2013). Natural Triple Excitations in Local Coupled Cluster Calculations with Pair Natural Orbitals. J. Chem. Phys. 139, 134101. doi:10.1063/1.4821834
- Schreier, W. J., Schrader, T. E., Koller, F. O., Gilch, P., Crespo-Herna'ndez, C. E., Swaminathan, V. N., et al. (2007). Thymine Dimerization in DNA Is an Ultrafast Photoreaction. *Science* 315, 625–629. doi:10.1126/ science.1135428
- Shaw, A. A., Falick, A. M., and Shetlar, M. D. (1992). Photoreactions of Thymine and Thymidine with N-Acetyltyrosine. *Biochemistry* 31, 10976–10983. doi:10.1021/bi00160a006
- Smith, K. C. (1962). Dose Dependent Decrease in Extractability of DNA from Bacteria Following Irradiation with Ultraviolet Light or with Visible Light Plus Dye. Biochem. Biophysical Res. Commun. 8, 157–163. doi:10.1016/0006-291X(62)90255-3
- Spitzer, J., Hafner, M., Landthaler, M., Ascano, M., Farazi, T., Wardle, G., et al. (2014). PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): A Step-by-step Protocol to the Transcriptome-wide Identification of Binding Sites of RNA-Binding Proteins. *Methods Enzymol.* 539, 113–161. doi:10.1016/B978-0-12-420120-0.00008-6
- Stingele, J., Bellelli, R., and Boulton, S. J. (2017). Mechanisms of DNA-Protein Crosslink Repair. Nat. Rev. Mol. Cel Biol 18, 563–573. doi:10.1038/nrm.2017.56
- Stützer, A., Welp, L. M., Raabe, M., Sachsenberg, T., Kappert, C., Wulf, A., et al. (2020). Analysis of Protein-DNA Interactions in Chromatin by UV Induced Cross-Linking and Mass Spectrometry. *Nat. Commun.* 11, 5250. doi:10.1038/ s41467-020-19047-7

- Sun, G., Fecko, C. J., Nicewonger, R. B., Webb, W. W., and Begley, T. P. (2006). DNA–Protein Cross-Linking: Model Systems for Pyrimidine–Aromatic Amino Acid Cross-Linking. Org. Lett. 8, 681–683. doi:10.1021/ol052876m
- Tretyakova, N. Y., Groehler, A., and Ji, S. (2015). DNA-protein Cross-Links: Formation, Structural Identities, and Biological Outcomes. Acc. Chem. Res. 48, 1631–1644. doi:10.1021/acs.accounts.5b00056
- Valadan, M., Pomarico, E., Della Ventura, B., Gesuele, F., Velotta, R., Amoresano, A., et al. (2019). A Multi-Scale Time-Resolved Study of Photoactivated Dynamics in 5-benzyl Uracil, a Model for DNA/protein Interactions. *Phys. Chem. Chem. Phys.* 21, 26301–26310. doi:10.1039/C9CP03839F
- Weigend, F. (2006). Accurate Coulomb-Fitting Basis Sets for H to Rn. Phys. Chem. Chem. Phys. 8, 1057–1065. doi:10.1039/B515623H
- Weigend, F., and Ahlrichs, R. (2005). Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* 7, 3297–3305. doi:10.1039/ B508541A
- Welsh, J., and Cantor, C. R. (1984). Protein-DNA Cross-Linking. Trends Biochem. Sci. 9, 505–508. doi:10.1016/0968-0004(84)90271-8
- Zhao, L., Zheng, H., Zhan, K., Guo, Y., Liu, B., and Xu, G. (2021). Position of the Benzene Ring Substituent Regulates the Excited-State Deactivation Process of the Benzyluracil Systems. J. Phys. Chem. A. 125, 165–174. doi:10.1021/ acs.jpca.0c08980
- Zheng, J., Xu, X., and Truhlar, D. G. (2011). Minimally Augmented Karlsruhe Basis Sets. *Theor. Chem. Acc.* 128, 295–305. doi:10.1007/s00214-010-0846-z
- Zhou, Q. Q., Zou, Y. Q., Lu, L. Q., and Xiao, W. J. (2019). Visible-Light-Induced Organic Photochemical Reactions through Energy-Transfer Pathways. *Angew. Chem. Int. Ed.* 58, 1586–1604. doi:10.1002/anie.201803102

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bende, Farcaş and Toşa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.