Eur. Phys. J. D (2017) 71: 155
DOI: 10.1140/epjd/e2017-70818-2

THE EUROPEAN PHYSICAL JOURNAL D

Regular Article

On binding specificity of (6–4) photolyase to a T(6–4)T DNA photoproduct[⋆,⋆⋆]

Katrine Aalbæk Jepsen and Ilia A. Solov'yov[a]

Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

Received 28 December 2016 / Received in final form 5 April 2017
Published online 15 June 2017 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2017

Abstract. Different factors lead to DNA damage and if it is not repaired in due time, the damaged DNA could initiate mutagenesis and cancer. To avoid this deadly scenario, specific enzymes can scavenge and repair the DNA, but the enzymes have to bind first to the damaged sites. We have investigated this binding for a specific enzyme called (6–4) photolyase, which is capable of repairing certain UV-induced damage in DNA. Through molecular dynamics simulations we describe the binding between photolyase and the DNA and reveal that several charged amino acid residues in the enzyme, such as arginines and lysines turn out to be important. Especially R421 is crucial, as it keeps the DNA strands at the damaged site inside the repair pocket of the enzyme separated. DNA photolyase is structurally highly homologous to a protein called cryptochrome. Both proteins are biologically activated similarly, namely through flavin co-factor photoexcitation. It is, however, striking that cryptochrome cannot repair UV-damaged DNA. The present investigation allowed us to conclude on the small but, apparently, critical differences between photolyase and cryptochrome. The performed analysis gives insight into important factors that govern the binding of UV-damaged DNA and reveal why cryptochrome cannot have this functionality.

1 Introduction

In everyday life biological organisms are exposed to different kinds of radiation. The radiation can be divided as either: ionizing radiation or non-ionizing radiation, depending on the energy of the radiation particles. Ionizing radiation has the energy exceeding 10 eV, which is high enough to ionize atoms and molecules, and break chemical bonds. The ionizing radiation could be produced by radioactive materials that emit α, β or γ radiation, other sources are X-rays and UV-light. The non-ionizing radiation covers the longer-wavelength of the electromagnetic spectrum and this radiation does not typically possess sufficient energy to ionize atoms or directly break chemical bonds.

One example of radiation damage could possibly occur in DNA after its exposure to ultraviolet (UV) radiation, for example from the sun. This could lead to the damage of the chemical structure of the double helix and lead to formation of a photoproduct between two adjacent pyrimidine rings in the DNA molecule. The possible photoproducts include a cyclobutane pyrimidine dimer (CPD) and the pyrimidine-pyrimidone (6–4) photoproduct [1], which are schematically illustrated in Figure 1.

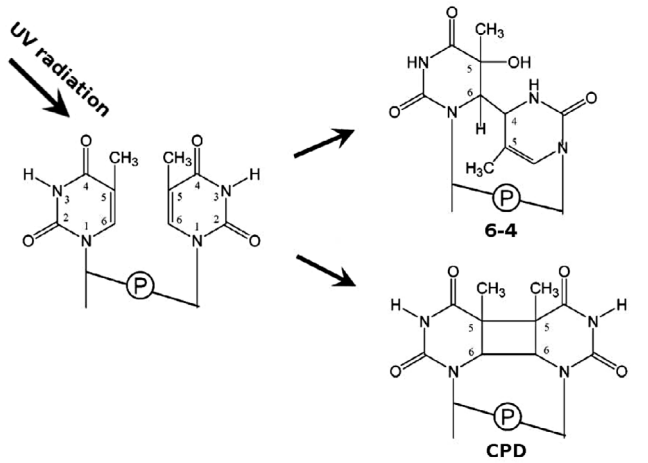

Fig. 1. UV induced photoproducts. UV radiation could introduce photoproducts between to adjacent pyrimidine rings in the DNA molecule. The most common photoproducts are cyclobutane pyrimidine dimer (CPD) and pyrimidine-pyrimidone (6–4) photoproduct. The photoproducts arising between two thymine base pairs (T-T) are shown, being the primary focus of this investigation.

If the photoproducts, or as they are also called photolesions, are not repaired in due time, the damaged DNA could cause mutagenesis and cancer [4]. The cellular machinery has, however, developed through evolution numerous biophysical mechanisms tailored to fight harmful perturbations [5]. To avoid the deadly scenario, a specialised

⋆ Supplementary material in the form of one pdf file available from the Journal web page at https://doi.org/10.1140/epjd/e2017-70818-2

⋆⋆ Contribution to the Topical Issue: "Dynamics of Systems at the Nanoscale", edited by Andrey Solov'yov and Andrei Korol.

[a] e-mail: ilia@sdu.dk

enzyme called as DNA photolyase was designed inside cells to interact with the photoproduct and efficiently repair it [6].

DNA photolyases are DNA UV-damage repair enzymes, consisting of 450–550 amino acid residues, having two cofactors bound internally [1], namely the flavin adenine dinucleotide (FAD) and either methenyltetrahydrofolate (MTHF) or 8-hydroxy-7,8-dede-methyl-5-deazariboflavin (8-HDF) [7]. FAD is an important catalyst in the DNA repair process, while the second cofactor is believed to serve as a photo antenna, which increases the repair rate 10–100 fold under dark conditions [1].

DNA photolesions could be repaired using two different types of DNA photolyase, which are called as the CPD photolyase and the (6–4) photolyase [1]. These two photolyases are damage specific, since CPD photolyase could only repair CPD photolesions, while the (6–4) photolyase could only fix the (6–4) photolesions. The CPD repair mechanism has been examined extensively during the past 30 years and is believed to be well characterised [3,8,9]. For the (6–4) repair mechanism several hypotheses have been proposed and the repair reaction is not as well understood as for the CPD case [3,10]. The repair of the (6–4) photoproduct is also seemingly more complicated than in the case of CPD, because an alcohol group is transferred from one nucleic base to the another upon DNA damage formation. Although the details of specific repair reactions are still debated, it is widely accepted [1,10–12] that in all photolyases the FAD cofactor is the primary cofactor in the repair process and prior the actual repair it is found in the two-electron-reduced $FADH^-$ form [1,3,9].

One of the proposed mechanisms of the (6–4) damage repair is illustrated in Figure 2B. First the active $FADH^-$ makes an electron transfer to the photoproduct. The electron could then be either futile back transferred or the photoproduct is repaired before this electron back transfer is facilitated. Another mechanism involves additional proton transfer from the H365 residue to the photoproduct, which accompanies the electron transfer [3].

In order for photolyase to repair the damaged DNA it has first to bind to the DNA photoproduct and then have the FAD cofactor in the active $FADH^-$ form. Once photolyase recognises a UV-damaged DNA it opens the DNA at the side of the lesion, and the lesion flips out into a hydrophobic pocket in the active site of the protein [2]. The FAD cofactor is found at the edge of this pocket. The bulge formed after flipping the damaged residues is stabilised by protein residues, which also prevent it from flipping back before the repair process is finalised.

Atomic information about the binding between the damaged DNA and photolyase is obtained from experimental crystal structure [2], which is static and does not deliver any dynamical aspects of the system that may arise in reality. It is also not possible to conclude on the binding energy between photolyase and the photoproduct and to justify which residues of the protein are most essential in the binding of the DNA. To investigate the dynamical features of the system, we have employed molecular dynamics (MD) simulations and addressed some basic questions on

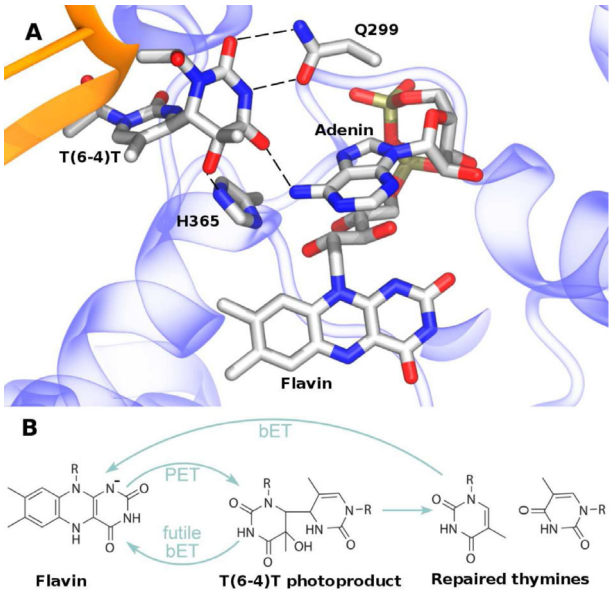

Fig. 2. Proposed mechanism for the (6–4) photo product repair. (A) The important residues from the (6–4) photolyase [2], suggested to be involved in the repair of the (6–4) DNA photoproduct. Q299 is important for keeping the photoproduct in place through hydrogen bonds [2], flavin is the primary electron donor [1] and H365 was suggested to be involved in a proton transfer in some mechanisms [3]. The dashed lines shows hydrogen bonds between residues. (B) The schematic repair mechanism of the T(6–4)T photoproduct where the fully reduced $FADH^-$ makes an electron transfer to the DNA photoproduct. The electron can either be futile back transferred or the photoproduct is repaired.

DNA-protein binding. From the performed MD simulation we obtain dynamic information about the binding site, and in particular establish the interaction energy between photolyase and the UV-damaged DNA. To elucidate the specifics of photolyase binding to a photolesion, we have considered the possible interaction of the UV-damaged DNA and the photolyase structural homology protein – cryptochrome. Cryptochromes occupy an important niche of the flavoprotein family; as photolyase they are activated by blue light through an internally bound FAD cofactor and the two proteins have a common ancestor with is a (6–4) photoproduct with a iron-sulfur cluster [13]. Despite the high sequence homology, see Figure 3, most cryptochromes are not able to repair photolesions [14], instead they are known to be involved in maintaining the circadian clock of a cell [1] and maybe endowing migratory birds with the magnetic compass sense [15–25]. Here, we compare the dynamics and the interactions energies of the two proteins, seeking to reason why cryptochromes are incapable of UV-damaged DNA repair, which photolyase does.

2 Methods

The binding of the UV-damaged DNA containing the T(6–4)T photoproduct to photolyase and cryptochrome

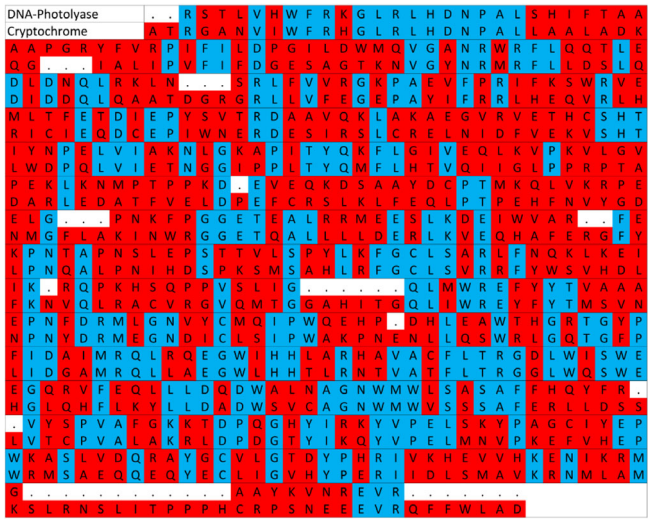

Fig. 3. Sequence similarity of *Drosophila melanogaster* photolyase and cryptochrome. Sequence comparison of the two structurally aligned proteins. The blue color indicates identical residues in both proteins, while the red color shows the different residues. The white color highlights gaps in the sequences of the two proteins.

was investigated using MD simulations. The simulations were performed through employing the NAMD 2.11 program [26] using the CHARMM22 force field with CMAP corrections for proteins [27,28]. For FAD and the T(6–4)T photoproduct a supplementary force fields were employed [19,20,29]. The supplementary force field parameters can be found in Supporting Material (SM). Analysis of the simulations results and visualisation of molecular structures were accomplished with VMD 1.9.2 [30].

2.1 Photolyase/DNA system preparation

The photolyase/DNA complex was constructed from the X-ray crystal structure of the protein bound to a double stranded DNA with 15 base pairs and the T(6–4)T photoproduct. The structure corresponded to *Drosophila melanogaster*, and was taken from the protein data bank (PDB ID: 3CVU) [2]. The missing hydrogen atoms in the X-ray structure were added to the system, which was then solvated within a water box modelled through the TIP3P force field [31]. The resulting solvated system was ionized with the NaCl salt concentration of 0.05 mol/L. The total number of atoms for the DNA/photolyase system result in 118 264, with the dimensions of the simulation box being equal to 110 Å × 120 Å × 98 Å.

In the simulations we do not consider the active form $FADH^-$ because this fully reduced form emerges after light induced electron transfer [11,12,29], and the binding of the DNA should not be affected by this event. There are still some controversies [3,11,12] regarding the activation mechanism of the 6–4 photolyase and, therefore, we believe that the resting state of FAD would be the optimal choice to address the question of the present investigation.

2.2 Cryptochrome/DNA system preparation

The X-ray crystal structure of cryptochrome from *Drosophila Melanogaster* was taken from the protein data bank (PDB ID: 4GU5) [32,33], and used here to construct the possible cryptochrome/DNA complex. There is no crystal structure of cryptochrome and a DNA with a T(6–4)T photoproduct available, so the UV-damaged DNA was added manually by first aligning the amino acids sequence of cryptochrome and photolyase, and then combining the structure of cryptochrome and the structure of the damaged DNA from the photolyase/DNA system. Due to the differences in cryptochrome and photolyase structures, a steric clash between the DNA and the protein was observed after the complex was constructed. This steric clash was handled by slightly moving the residues V297, R298 and G299 away from the DNA attached to the structure of cryptochrome. The missing hydrogen atoms were added to the system, and the system was solvated and ionized in the same way as the photolyase/DNA system. The solvated cryptochrome/DNA complex resulted in having 120 286 atoms in the simulation box of 189 Å × 127 Å × 91 Å.

2.3 Molecular dynamics simulations

All MD simulations were performed with the 2 fs time step. The cut off distances for the van der Waals and Coulomb interactions were set to 12 Å, where the long-range electrostatic interactions were treated using the PME method [34], employing periodic boundary conditions. The NPT ensemble was used for the equilibration, with a temperature of 310 K by applying Langevin forces with a damping coefficient of 1 ps^{-1}. The pressure control was achieved through the Nosé-Anderson-Langevin piston [35], keeping the average pressure at a value of 1 atm. The piston oscillation period was set to 100 fs and the damping scale was put equal to 50 fs. Pressure control was used for the equilibration simulations to ensure setting of the correct density of the system and the production simulation employed the NVT ensemble. The complete simulation protocol, is summarised in Table 1.

Equilibration of the photolyase/DNA complex was performed after the initial energy minimization of the system. In the first stage of equilibration the atoms in the protein, the FAD co-factor and the DNA were harmonically constrained during 1 ns interval, which permitted the water and the ions to relax. Afterwards the side chains of the protein and the DNA were allowed to move, while the simulation was carried out for another 1 ns. The third stage of the equilibration involves all atoms in the system to be released and the system was further simulated for 5 ns, still in the NPT ensemble. The statistical ensemble was changed to NVT and the system was simulated for equilibration purposes for another 150 ns, which was followed by a 150 ns production simulation.

The protocol was slightly different in the case of the cryptochrome/DNA simulation. Because of the residues

Table 1. Simulation protocol employed in the investigation. Both, photolyase/DNA and cryptochrome/DNA, systems consist of the protein, FAD co-factor and the UV-damaged DNA double strand with a T(6–4)T photoproduct. The systems were first equilibrated in the NPT ensemble where the atoms in the systems were initially constrained, and the constrains were gradually removed before the production simulations started, which were carried out in the NVT ensemble.

Photolyase/DNA		Cryptochrome/DNA	
Structure minimization			
5000 NAMD time steps			
Process	Simulation time (ns)	Process	Simulation time (ns)
Equilibration			
Water box equilibration, rest constrained	1	Residues V297, R298 and G299 and water box equilibrium, rest constrained	0.5
Side chains released for protein and DNA	1	DNA side chains released	2
All atoms released	5	DNA released	1
		Protein side chain released	1
		All atoms released	5
Production			
Equilibration in NVT ensemble	150	Equilibration in NVT ensemble	100
Production MD run	150	Production MD run	150

that clashed with the DNA after attaching DNA to cryptochrome (residues V297, R298 and G299) the equilibration of the cryptochrome/DNA complex was performed differently. Equilibration of the three residues was initially performed, while harmonic constrains were applied to the protein and the DNA for 0.5 ns with 1 fs time step. Afterwards the side chains of the DNA were released and simulated for further 2 ns, before the rest of the DNA was released and simulated for 1 ns. The next equilibrium phase involves protein equilibration: first for 1 ns with constrained backbone and then with all atoms released for another 5 ns before the statistical ensemble was changed to NVT, and the entire system was simulated for another 100 ns, followed by a 150 ns production simulation.

3 Results

To investigate the binding between the UV-damaged DNA and photolyase we first consider the stability of the protein upon binding to the DNA before we proceed with analysing the interaction energy and binding specificity of the protein and the DNA, which is further evidenced through the difference from the cryptochrome/DNA system.

3.1 Stability of photolyase-DNA binding

An important measure for molecular system stability analysis is the root mean square displacement (RMSD), which indicates how much the atoms in a molecular structure are displaced on average compared to a reference position.

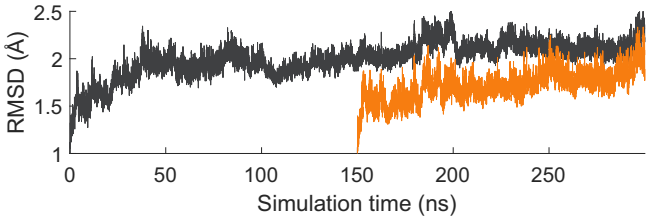

Fig. 4. Stabilising photolyase-DNA complex. Root mean square displacement (RMSD) computed for photolyase employing equation (1) relative to the structure obtained after the initial equilibration (black line) and relative to the structure obtained after 150 ns (orange line) of simulation.

RMSD is typically defined as:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N} \left(r_i\left(t_1\right) - r_i\left(t_2\right)\right)^2}{N}}, \qquad (1)$$

where N is the total number of atoms in the studied structure, and $r_i(t_1)$ and $r_i(t_2)$ denote positions of an atom i at time instances t_1 and t_2, respectively.

The computed RMSD time evolution for photolyase after all atoms were released in the simulation, see Table 1, is shown in Figure 4. The plot indicates that the RMSD value of the protein showed saturation after about 150 ns. The first 150 ns of the simulation where thus not used for follow up any analysis and the simulation continued for another 150 ns. The orange line shows the time evolution of RMSD compared relatively to the structure obtained after 150 ns simulation. The plot shows that choosing a reference structure at a later time instance leads to the RMSD decrease and indicates that there are no significant structural displacements in the protein. The orange line

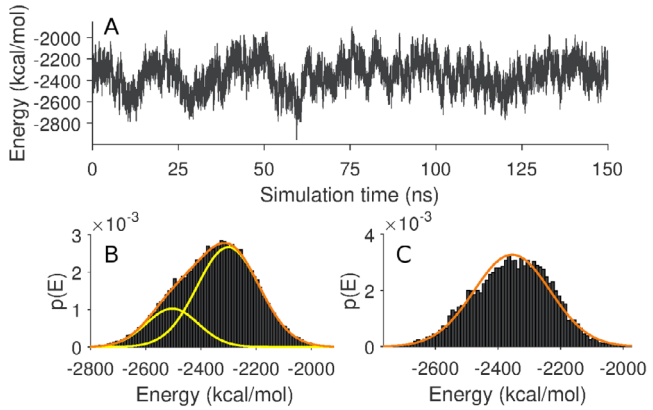

Fig. 5. DNA binding to photolyase. (A) Time evolution of the interaction energy between the UV-damaged DNA and photolyase. (B) Probability density distribution computed for the interaction energy between DNA and photolyase, and the corresponding Gaussian fit. (C) Probability density distribution computed for the interaction energy between DNA and photolyase during the last 50 ns of the simulation, and the corresponding Gaussian fit.

in Figure 4 shows little increase in value, but it is in order of no significant importance as a RMSD value of 1.5–2 Å is considered small for a system of this size [24].

3.2 Binding of the UV-damaged DNA to photolyase

The binding strength of the UV-damaged DNA to photolyase could be characterised through the interaction energy shown in Figure 5. The interaction energy between the protein and the damaged DNA molecule accounts for the two non-bonded contributions, namely van der Waals and electrostatic interactions. The average interaction energy between the studied fragment of the damaged DNA and the entire photolyase is −2351.1 kcal/mol, where −129.7 kcal/mol are due to van der Waals interactions and the remaining −2221.4 kcal/mol are of electrostatic nature.

In this context it is useful to consider the results of the present investigation from the perspective of the ergodic hypothesis, which suggest that the average of a process parameter over time and the average over the statistical ensemble should be the same. We can, therefore, use the time evolution of a process parameter to find the average over the statistical ensemble. The hypothesis can be applied to energy calculations delivered from MD simulations. In this case the probability density distribution of the interaction energy is expected to be fitted with a single Gaussian distribution, being defined as:

$$p(E) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(E - \langle E\rangle)^2}{2\sigma^2}\right), \qquad (2)$$

where $\langle E\rangle$ is the mean energy value, σ is the standard deviation and σ^2 is the energy variance.

The interaction energy of the DNA and photolyase, computed in the course of MD simulation, is used to compute the probability density distribution of the interaction energy, as shown in Figures 5B and 5C for two different time intervals taken from the simulation. The probability density distribution of the interaction energy between the damaged DNA and photolyase in Figure 5B features an asymmetric distribution, which could be fitted with two Gaussians, indicating that the damaged DNA might bind to photolyase following two different structural motifs; each structural binding motif corresponds to one Gaussian distribution. The time evolution of the interaction energy in Figure 5A shows that the energy oscillates between the two states in the beginning of the 150 ns production simulation, but later converges to a more steady binding regime, characterised by the average interaction energy of –2301.15 kcal/mol. The oscillating behaviour in the beginning of the simulation indicates that the system has not found a stable conformation, and continuously switches between two different binding modes. The probability density distribution of the interaction energy computed for the last 50 ns of the simulation trajectory in Figure 5C is nicely fitted with a single Gaussian profile. The shifts in the binding energy can be attributed to changes in the DNA conformation.

The performed analysis indicates that it takes the 150 ns equilibration plus extra 75 ns before a stable binding behaviour of the DNA molecule to the repair enzyme is revealed. The increased simulation time, could arise due to the presence of a DNA in the system. The DNA used in this simulation has only 15 base pairs, which makes its ends easily move around and interact with the protein. Simulation indicate a stable binding during the final 50 ns, of the simulation but this high DNA flexibility could in principle make the DNA readjust its position if the simulation time would be increased.

For the repair process to function properly, it is advantageous that the interaction energy between the UV-damaged DNA and photolyase is negative as it reveals attraction between the fragments. However, it is even more important for the repair process to have negative interaction energy between photolyase and the DNA photoproduct as it shows binding specificity. Figure 6A shows this component of the interaction energy computed in the course of MD simulation with the average value of –236.5 kcal/mol. The interaction energy between the photoproduct and photolyase follows, a trend similar to the entire interaction energy, the double peaked probability density distribution, as shown in Figure 6B. In the case of this specific binding interactions energy, the second broad peak arises due to somewhat increased binding energy during the first 50 ns of the simulation, which then stabilises around an average value on –248.41 kcal/mol. Similar to Figure 5C the probability density distribution of the interaction energy computed between the photoproduct and photolyase for the last 50 ns of the simulation can be fitted with a single Gaussian profile, as featured in Figure 6C.

Even though the total interaction energy between DNA and photolyase changes due to the high flexibility

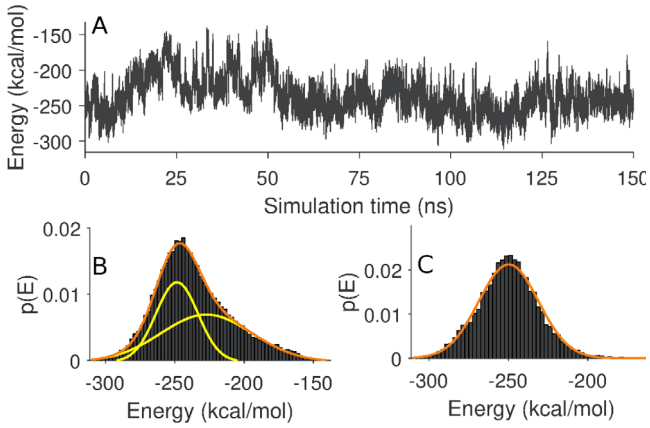

Fig. 6. The T(6–4)T photoproduct binding to photolyase. The interaction energy between the T(6–4)T photoproduct and photolyase where A shows the time evolution, B shows the probability density distribution, and C shows the probability density distribution for the last 50 ns of the simulation.

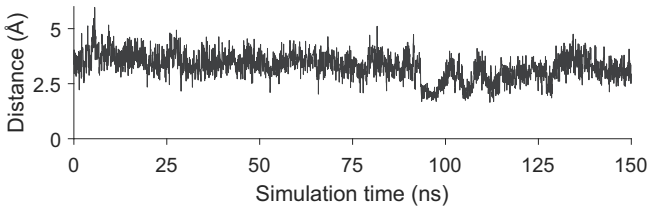

Fig. 7. Distance between the FAD and the photoproduct. The time evolution of the edge-to-edge distance between the FAD cofactor and the photoproduct in the UV-damaged DNA.

of the DNA, it seems like this effect does not affect the ability of photolyase to repair the photoproduct. Because the repair mechanism involves an electron transfer from FAD to the photoproduct [2,10,36], an increased distance between these two components will affect the probability of the repair process [29]. As seen in Figure 7, the edge-to-edge distance between the FAD cofactor from photolyase and the photoproduct is stable at around 3.5 Å most of the time, and there is no noticeable increases over time. The only peculiarity is observed at 100 ns, which could be connected to DNA internal dynamics. At this time instance the ends of the DNA shortly move away from the protein decreasing the angle of the backbone at the site of the photoproduct, making it possible for the photoproduct to move closer to the FAD.

3.3 Important factors for DNA binding

The performed simulations permit revealing the specific role of individual residues that participate in DNA binding to photolyase. Table 2 summarises the most important residues, lists how often the residues form hydrogen bonds and compiles the mean interaction energies of these residues with the DNA.

The residues with the lowest value of the interaction energy are the positively charged amino acids lysines and arginines. These amino acids bind to the negatively charged backbone of the DNA and keep it in place. Figure 8A shows photolyase colored after the electrostatic properties of amino acids; the positively charged amino acids are shown in blue and the negatively charged are shown in red. It is worth noting that at the site where the DNA backbone approaches the protein, the surface is blue and thereby enhancing the binding of the negatively charged DNA.

The binding of the photoproduct is strongly relying on two key residues with comparable interaction energies, namely K246 and R421, see Figure 9. The interaction energy of –86.03 kcal/mol for K246 and the photoproduct is the lowest value among all residues in the protein and in the DNA, see Figure 10. The structural analysis made of Maul et al. [2] also suggest that the K246 residue is key for photolyase binding. During the simulation K246 moves from interacting with the end of the photoproduct to interact with the negative phosphate group at the backbone between the two nucleotides forming the photoproduct. This shift can be connected to the change in interaction energy seen in Figure 6. Analysing the crystal structure suggest that K246 points towards the FAD, and one could expect that it would interact more with the FAD than with the photoproduct. The other residue, R421, on the other hand is already known to be an important factor in the binding of UV-damaged DNA to photolyase [2]. The time average interaction energy of R421 with the photoproduct equals −55.36 kcal/mol, while it is −75.28 kcal/mol for the interaction energy between R421 and base 10 on DNA strand C, suggesting that R421 could prevent the DNA strands from reconnecting but also stabilises the binding of photolyase and DNA. Some other residues important to keep the DNA in place are K488, R502 and R505, as can be seen from Figure 10. With interactions energies between –25.32 and –66.00 kcal/mol these residues are noticeably contributing to stability of photolyase/DNA attachment. The three residues are placed on the same α-helix, see Figure 9, and all interact with the DNA strand D.

3.4 Protonation of H365

It has been proposed earlier [2], that the binding of the photoproduct is additionally mediated by the H365 residue from photolyase, as it forms hydrogen bonds with the photoproduct. In the performed simulation no long-term hydrogen bonds were observed. The simulations indicate that H365 rotates early in the simulation and, therefore, it is not possible for this residue to make the predicted hydrogen bonding network. This situation can be a consequence of the decided protonation state of the histidine; H365 was assigned to be protonated in the δ-position. If the histidine was protonated on the ϵ-position or on both nitrogen positions, it is likely that the histidine would behave differently.

The change in H365's position has also an effect on the photoproduct: in the crystal structure it seems like it can form, two hydrogen bonds with Q299, see Figure 2, but in the simulation one could only identify a transient single

Table 2. Hydrogen bonds between the UV-damaged DNA and photolyase. The donor and acceptor are the residues which build the hydrogen bonds investigated. The donor residue contains the hydrogen covalently bound to the electronegative atom involved in the hydrogen bond, while the acceptor contains the electronegative atom not covalently bound to a hydrogen bond. The C and D denote the two DNA strands, such that the chain C indicates the DNA strand containing the photoproduct, while strand D represents the other strand. The interaction energy represents the sum of the van der Waals and electrostatic interactions involved in the corresponding hydrogen bond, sorted by increasing values. The occupancy is the percent fraction of the simulation time when the donor and acceptor are at a distance from each other where they are capable of forming a hydrogen bond.

Donor	DNA strand	Acceptor	Occupancy	Mean interaction energy (kcal/mol)
K246	C	T(6–4)T	57.58%	–86.03
R421	C	G10	34.59%	–75.28
R502	D	C10	21.16%	–66.00
K161	D	T14	61.68%	–58.51
K431	C	C11	76.31%	–58.26
R502	D	C11	57.80%	–58.12
R421	C	T(6–4)T	4.17%	–55.36
K154	D	C13	7.10%	–43.91
R505	D	C11	72.28%	–43.32
K488	D	A8	11.30%	–25.32
Y419	D	C10	68.21%	–13.12
S424	C	C11	81.11%	–11.91
Q299	C	T(6–4)T	4.97%	–5.02

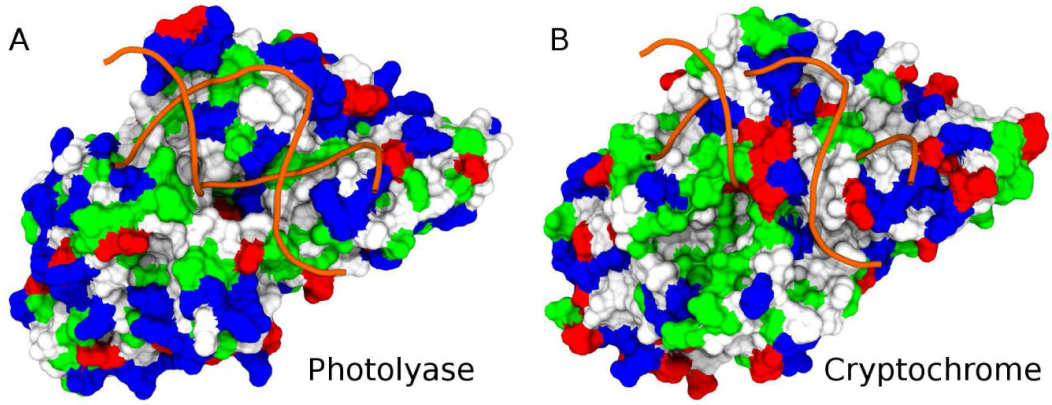

Fig. 8. Photolyase and cryptochrome. The proteins are colored after amino acid type, such that, blue shows basic, red shows acidic, green shows polar and white the nonpolar residues. Both DNA's are taken from the photolyase crystal structure [2] and placed in cryptochrome after structural alignment of the two proteins, see methods.

hydrogen bond for about 4.97% of the entire simulation time.

3.5 Cryptochrome binding to UV-damaged DNA

Photolyase and cryptochrome have high sequence similarity [1], but their biological functions are apparently completely different [1]. These differences could partially stem from the fact that cryptochrome does not have a favourable binding mode with the UV-damaged DNA.

The average interaction energy between the damaged DNA and photolyase is −2349.24 kcal/mol, while it is +64.14 kcal/mol for the damaged DNA and cryptochrome, see Figure 11A. Because of the large difference it is clear that photolyase would bind to the DNA significantly better than cryptochrome. The positive value for cryptochrome indicates that the DNA binds weakly and it actually is even expected to dissociate after some time, however the performed 150 ns simulations did not reveal this. A plot demonstrating the time evolution of the distance between cryptochrome and DNA is given in Supporting Material (SM) in Figure S1. The fact that cryptochrome does not unbind from the DNA could be an effect from the negative interaction energy between the protein and the photoproduct. Cryptochrome binds the photoproduct with an energy on –66 kcal/mol, while photolyase binds the photoproduct with an average interaction energy on –230 kcal/mol, see Figure 11B. As expected photolyase binds stronger to the photoproduct than cryptochrome.

When comparing the surface of photolyase and cryptochrome, see Figure 8, none of the surface around the DNA binding site in photolyase is showing acidic behavior (colored red). For cryptochrome, on the other hand, the acidic residue E528 is placed at the position, where

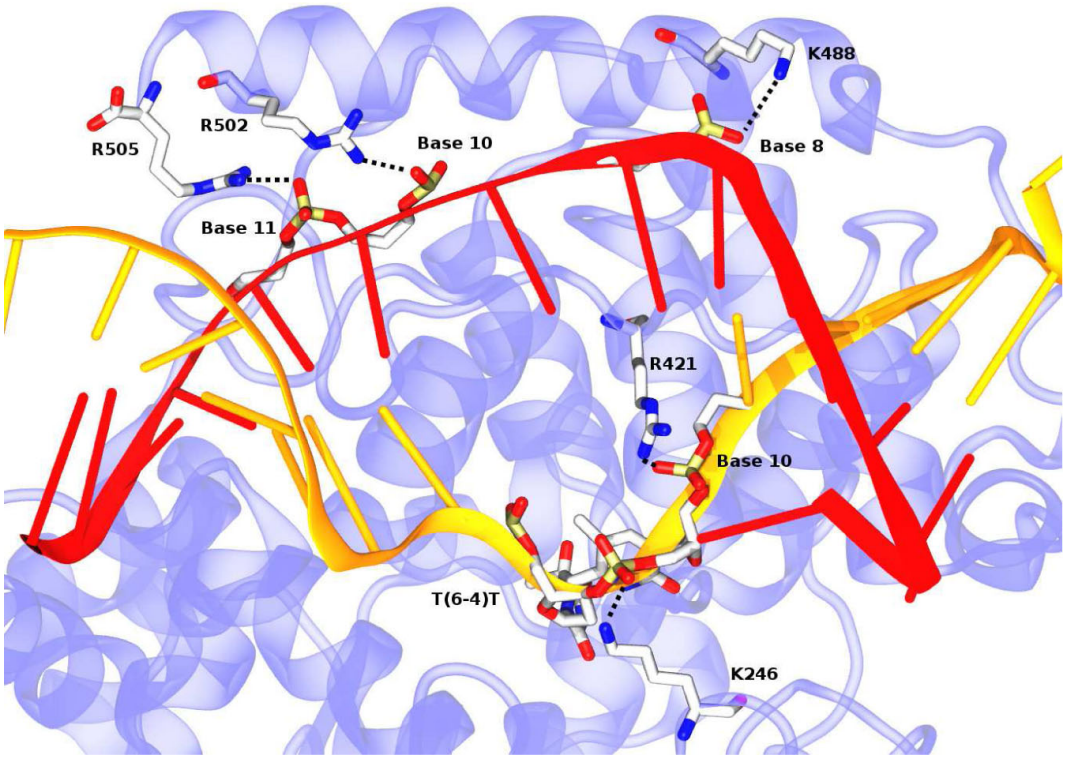

Fig. 9. Important photolyase residues for DNA binding. Residues K488, R502 and R505 form hydrogen bonds with the DNA base residues 8, 10 and 11 from the DNA-D strand, while the residues K246 and R521 form hydrogen bonds with the photoproduct and DNA base residue 10 from DNA-C strand.

Table 3. Residues involved in DNA binding to photolyase as compared with cryptochrome. The photolyase residues are highlighted here, being the most important amino acids for the binding of the UV-damaged DNA. The equivalent residues in cryptochrome are those found at the same position as the residues in photolyase after structural alignment of the two proteins. The third column shows the mean interaction energy (in kcal/mol) between a residue in cryptochrome and the closest DNA residue.

Residue in photolyase	Equivalent residue in cryptochrome	Interaction energy with the nearest DNA residue (kcal/mol)
R505	P520	0.35
R502	I517	0.12
K488	K503	–10.16
Y419	–	
K161	M160	0.32
K431	R446	–33.74
S424	C439	–1.34
R421	S434	–1.55
K246	–	

DNA is supposed to bind; the negative charge on the glutamic acid will repeal the negatively charged DNA. Other differences between photolyase and cryptochrome are the properties of the individual amino acid residues in cryptochrome located at the same locations as the residues from Table 2 in photolyase; the differences are highlighted in Table 3. In most of the cases the residues in cryptochrome are non-polar and neutral, making them weaker binding partners for the negatively charged DNA. Only in the case of K431 and K488 in photolyase the residue in cryptochrome is also a positively charged amino acid residue. Of the residues listed in Table 3 the two positive charged residues are the only having a interaction energy comparable with the same residues in photolyase. The non-polar and neutral residues, together with the negatively charged E528 makes it difficult for cryptochrome to bind the UV-damaged DNA. Therefore, the DNA has moved a little bit away from the proposed binding site during the simulation, in order to bind to cryptochrome. Here it interact with the positively charged amino acid residue R532 in a way resembling the binding between R421 and the UV-damaged DNA in photolyase. They are both located between the two DNA strands and the energy between the residues and the photoproduct are –77.73 kcal/mol for cryptochrome and –49.52 kcal/mol for photolyase, respectively.

4 Conclusion

The performed simulations reveal that the binding between the UV-damaged DNA and photolyase is primarily governed by the electrostatic interactions between the positive charges on the protein surface and the negatively charged backbone of the DNA. In contrast, the absence of the positively charged amino acids arginine and

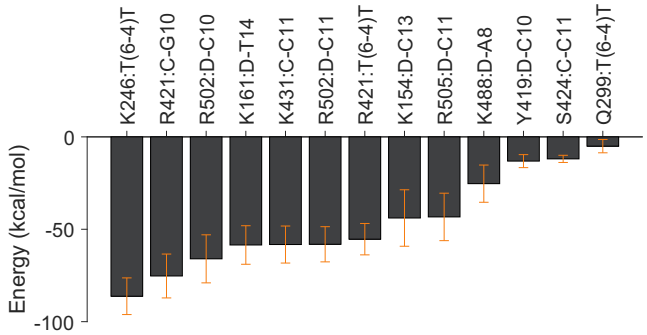

Fig. 10. Interactions between photolyase key residues and the UV-damaged DNA. Graphical illustration of the average interaction energies between amino acid residues in photolyase and residues of the UV-damaged DNA. The standard deviations of the energies are shown as error bars. The figure follows data compiled in Table 2.

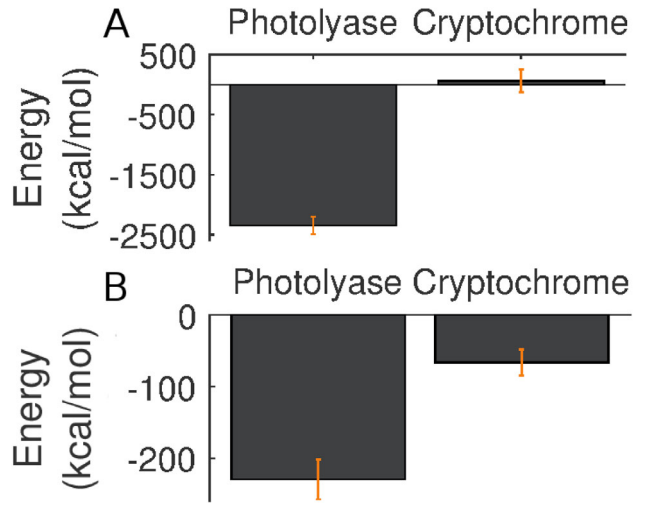

Fig. 11. Comparison of photolyase and cryptochrome binding to the UV-damaged DNA. The plot shows the mean energy and the standard deviation for the interaction energy between the DNA and either photolyase or cryptochrome (A). The interaction between the photoproduct and both proteins (B) is also presented.

lysine together with the presence of the negatively charged glutamic acid residue in cryptochrome makes the cryptochrome/DNA binding difficult.

The UV-damaged DNA is a highly flexible molecule, and the present investigation showed that significant equilibration time is required before a stable binding motif of photolyase to DNA could be observed. Such a behaviour is important to consider in the description of the complete repair process as it is a necessity for an efficient electron transfer between the repair enzyme and the damaged DNA site.

The repair of the (6–4) photoproduct by photolyase is one of the simple examples where DNA damage arises. Radiation that is more intense than UV radiation can lead to single and double strand breaks, and the repair of these damages are more complicated and their investigation is more involved. The present study could, however, be useful for such future investigations as it describes accurately and systematically how to characterise protein binding to a damaged DNA site. An interesting example could be ligase III, involved in repair of single strand breaks [37,38], and, therefore, such an investigation would require novel computational approaches that permit studying chemical bond breaking and formation which could, for example, be modeled with the MBN Explorer software [39], a universal computational tool tailored for multiscale simulations.

5 Supporting material

The force field parameters for the T(6–4)T photoproduct are provided in the form of CHARMM topology and parameter files in the Supporting Material. A Supplementary Figure S1 featuring the time evolution of the distance between cryptochrome and DNA is also included.

The authors thank the Lundbeck Foundation and the Russian Science Foundation (Grant No. 14-12-00342) for financial support. Computational resources for the simulations were provided by the DeIC National HPC Center, SDU.

Author contribution statement

I.A.S. designed the study. K.A.J. conducted calculations and performed analysis. I.A.S. and K.A.J. wrote the manuscript and prepared the figures.

References

1. A. Sancar, Chem. Rev. **103**, 2203 (2003)
2. M.J. Maul, T.R.M. Barends, A.F. Glas, M.J. Cryle, T. Domratcheva, S. Schneider, I. Schlichting, T. Carell, Angew. Chem. Int. Ed. **47**, 10076 (2008)
3. Z. Liu, L. Wang, D. Zhong, Phys. Chem. Chem. Phys. **17**, 11933 (2015)
4. D.E. Brash, Trends Genet. **13**, 410 (1997)
5. A. Sancar, L.A. Lindsey-Boltz, K. Ansal Kaçmaz, S. Linn, Annu. Rev. Biochem. **73**, 39 (2004)
6. P. Heelis, S.T. Kim, T. Okamura, A. Sancar, J. Photochem. Photobiol. B **17**, 219 (1993)
7. A. Sancar, Biochemistry **33**, 2 (1994)
8. Y.T. Kao, C. Saxena, L. Wang, A. Sancar, D. Zhong, Proc. Natl. Acad. Sci. **102**, 16128 (2005)
9. K. Brettel, M. Byrdin, Curr. Opin. Struct. Biol. **20**, 693 (2010)
10. T. Domratcheva, J. Am. Chem. Soc. **133**, 18172 (2011)
11. F. Cailliez, P. Müller, T. Firmino, P. Pernot, A. de la Lande, J. Am. Chem. Soc. **138**, 1904 (2016)
12. I.M.M. Wijaya, T. Domratcheva, T. Iwata, E.D. Getzoff, H. Kandori, J. Am. Chem. Soc. **138**, 4368 (2016)
13. F. Zhang, P. Scheerer, I. Oberpichler, T. Lamparter, N. Krauss, Proc. Natl. Acad. Sci. **110**, 7217 (2013)
14. A. von Zadow, E. Ignatz, R. Pokorny, L.O. Essen, G. Klug, FEBS J. **283**, 4291 (2016)

15. D.R. Kattnig, I.A. Solov'yov, P.J. Hore, Phys. Chem. Chem. Phys. **70**, 12443 (2016)
16. D.R. Kattnig, J.K. Sowa, I.A. Solov'yov, P.J. Hore, New J. Phys. **18**, 063007 (2016)
17. A. Möller, S. Sagasser, W. Wiltschko, B. Schierwater, Naturwissenschaften **91**, 585 (2004)
18. M. Liedvogel, H. Mouritsen, J. R. Soc. Interface **7**, S147 (2010)
19. G. Lüdemann, I.A. Solov'yov, T. Kubař, M. Elstner, J. Am. Chem. Soc. **137**, 1147 (2015)
20. I.A. Solov'yov, T. Domratcheva, K. Schulten, Sci. Rep. **4**, 3845 (2014)
21. I.A. Solov'yov, K. Schulten, J. Phys. Chem. B **116**, 1089 (2012)
22. I.A. Solov'yov, D.E. Chandler, K. Schulten, Biophys. J. **92**, 2711 (2007)
23. I. Solov'yov, K. Schulten, Biophys. J. **96**, 4804 (2009)
24. I.A. Solov'yov, T. Domratcheva, A.R. Moughal Shahi, K. Schulten, J. Am. Chem. Soc. **134**, 18046 (2012)
25. I.A. Solov'yov, T. Ritz, K. Schulten, P.J. Hore, *Quantum Effects in Biology* (Cambridge University Press, 2014), Chap. 10, pp. 218–236
26. J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, J. Comput. Chem. **26**, 1781 (2005)
27. A.D. MacKerell Jr., M. Feig, C.L. Brooks III, J. Comput. Chem. **25**, 1400 (2004)
28. A.D. MacKerell Jr., D. Bashford, M. Bellott, R.L. Dunbrack Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha et al., J. Phys. Chem. B **102**, 3586 (1998)
29. E. Sjulstok, J.M.H. Olsen, I.A. Solov'yov, Sci. Rep. **5**, 18446 (2015)
30. W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. **14**, 33 (1996)
31. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. **79**, 926 (1983)
32. B. Zoltowski, A. Vaidya, D. Top, J. Widom, M. Young, B. Crane, Nature **480**, 396 (2011)
33. B. Zoltowski, A. Vaidya, D. Top, J. Widom, M. Young, B. Crane, Nature **496**, 252 (2013)
34. T. Darden, D. York, L. Pedersen, J. Chem. Phys. **98**, 10089 (1993)
35. S.E. Feller, Y.H. Zhang, R.W. Pastor, B.R. Brooks, J. Chem. Phys. **103**, 4613 (1995)
36. A.R. Moughal Shahi, T. Domratcheva, J. Chem. Theor. Comput. **9**, 4644 (2013)
37. K.W. Caldecott, Exp. Cell Res. **329**, 2 (2014)
38. E. Cotner-Gohara, I.K. Kim, M. Hammel, J.A. Tainer, A.E. Tomkinson, T. Ellenberger, Biochemistry **49**, 6165 (2010)
39. I.A. Solov'yov, A.V. Yakubovich, P.V. Nikolaev, I. Volkovets, A.V. Solov'yov, J. Comput. Chem. **33**, 2412 (2012)