

A QM/MM study of the initial excited state dynamics of green-absorbing Proteorhodopsin

Veniamin A. Borin,^a Christian Wiebeler^a and Igor Schapiro^a

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

The primary photochemical reaction of the green-absorbing Proteorhodopsin is studied by means of a hybrid quantum mechanics/molecular mechanics (QM/MM) approach. The simulations are based on a homology model derived from the blue-absorbing Proteorhodopsin crystal structure. The geometry of retinal and the surrounding sidechains in the protein binding pocket were optimized using the QM/MM method. Starting from this geometry the isomerization was studied with a relaxed scan along the C₁₃=C₁₄ dihedral. It revealed an “aborted bicycle pedal” mechanism of isomerization that was originally proposed by Warshel for bovine rhodopsin and bacteriorhodopsin. However, the isomerization involved the concerted rotation about C₁₃=C₁₄ and C₁₅=N, with the latter being highly twisted but not isomerized. Further, the simulation showed an increased steric interaction between the hydrogen at the C₁₄ of the isomerizing bond and the hydroxyl group at the neighbouring tyrosine 200. In addition, we have simulated a nonadiabatic trajectory which showed the timing of the isomerization. In the first 20 fs upon excitation the order of the conjugated double and single bonds is inverted, consecutively the C₁₃=C₁₄ rotation is activated for 200 fs until the S₁-S₀ transition is detected. However, the isomerization is reverted due to the specific interaction with the tyrosine as observed along the relaxed scan calculation. Our simulations indicate that the retinal - tyrosine 200 interaction plays an important role in the outcome of the photoisomerization.

Introduction

In 2000 Béjà et al.^{1,2} have discovered a protein in γ -proteobacteria that later turned out to be the most abundant member of the rhodopsin protein family.³ Rhodopsins are light-sensitive proteins that consist of the apoprotein opsin which is characterized by seven transmembrane helices and a prosthetic group retinal that serves as a chromophore. This chromophore is covalently bound to the apoprotein through a Schiff base bond to the lysine sidechain making it effectively a photoreceptor protein. This composition forms the structural basis for the members of the rhodopsin family that are absorbing light in the UV/Vis range. Upon light absorption, the retinal undergoes isomerization that represents the first step in a series of conformational changes summarized in a photocycle that ultimately leads to a molecular response. The exact function of this light-induced activation and the fine details are varying within the rhodopsin family. The family is distinguished in two groups:⁴ the vertebrate rhodopsins (type II) and microbial rhodopsins (type I). Type I rhodopsin starts from an all-*trans* retinal protonated Schiff base (rpSb) and has a closed photocycle, in which the retinal chromophore is not released from the protein and the reisomerization occurs thermally. Type II rhodopsin binds 11-*cis* rpSb and has an open photocycle,

where the retinal can be released from the protein and reisomerised by enzymes or via a photochemical reaction. PR is categorized as type I rhodopsin and is closely related to its well-studied subfamily-representative bacteriorhodopsin (bR).⁵

The overall sequence identity between PR and bR is less than 30%,¹ nevertheless most of the functional amino acids are conserved.^{6,7} Among them are R94 and D227 (corresponding to Arg82 and D212 in bR) that form the counterion complex. The primary proton acceptor is D97 in PR and D85 in bR while the primary proton donor is homologue Glu108 and D96, respectively. These amino acids are essential for the function of bR as a proton pump that translocates protons from the cytoplasmic to the extracellular side across the membrane. Hence, it is not surprising that PR was confirmed to be a proton pump as well.^{1,2,6}

PR has some characteristics that make it distinct from bR despite the similar functional role and the preserved key amino acids for the pumping. One of them is the unusually high pK_a value of 7.5 of the primary proton acceptor D97 (D82 in bR has a pK_a of 2.5) which means that it is protonated to a large degree at physiological conditions (pH 7-8). Since the primary proton acceptor is protonated, the proton transfer from the retinal pSB is delayed or even aborted at very low pH values (<2.5).⁸ That means PR is operating at alkaline pH values. The protonation state has also an influence on the photochemistry of PR. The deprotonation of D97 leads to a spectral shift from 544 nm at pH 5 to 516 nm at pH 9.⁹ This effect of the electrostatic environment on the wavelength or colour of the absorbed light by the chromophore can be explained by the point charge model.¹⁰ Briefly, the chromophore absorption is fine-tuned by

^a Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

the protein surrounding, in particular by the immediate environment in the so-called protein binding pocket. Due to the intramolecular charge-transfer of retinal after excitation, a differential effect can be achieved by introduction of amino acid mutations that will influence the ground state to a different extent than the excited state resulting in an overall spectral shift.¹¹ There are 23 amino acids which have at least one atom with 5 Å distance from the chromophore. Interestingly, one of these amino acids (position 105) is found to be responsible for colour tuning in PR. Therefore, the PR family is divided into two subgroups depending on whether this residue is a leucine or a glutamine. The PR subfamily with L105 absorbs green light (GPR), while the other group with Q105 absorbs blue light (BPR)¹². This is due to difference in the depth of the ocean where the bacteria were found. As blue light is penetrating deeper into water, the bacteria that are found at lower level of the ocean have adapted to absorb this wavelength of light.^{3,13} Comparison of the protein sequence lead to the finding that only one amino acid in the protein binding pocket is largely responsible for the spectral shift of 35 nm. A single point mutant L105Q recovered 25 nm of the shift, although 4 of the 23 amino acids in close proximity are different (F137, A151, F152, and C156 in GPR and G137, W151, W152, and M156 in BPR). In this sense, the proteorhodopsin serves as an ideal system to study the spectral tuning.

However, also the primary dynamics following the excitation have been of interest in order to see for example if the excited state properties can be affected by mutations.^{14–19} A relative shift between the ground and excited state potential energy surfaces could lead to changes in the excited state lifetimes or the quantum yield of the photoreaction. To study such ultrafast dynamics, time-resolved absorption and infrared spectroscopy measurements have been conducted mainly on GPR.^{20–26} The ultrafast dynamics of native GPR at pH 9 was found to be best described by an excited state relaxation and a biexponential decay with time constants of 140 fs, 280 fs and 9.5 ps, respectively.²¹ However, the underlying molecular mechanism of the isomerization has not been explored by computational methods, mainly due to the lack of a crystal structure of GPR.

In order to perform computational investigations of PRs, a structural model of the protein is required. So far, the crystal structure of GPR is not solved and only recently, this was achieved for BPR.²⁷ Therefore, previous studies have used homology models of BPR and GPR based on the crystal structures of bR or SRII as a template.^{28,29} Excited state properties have been determined based on a semiempirical method by employing cluster models (rpSb and a few adjacent amino acids) from classical molecular dynamics (MD) simulations. Furthermore, the crystal structure of BPR served as a starting point to investigate ground state properties via classical MD.³⁰ Finally, Buda *et al.*³¹ have also used cluster models to calculate the properties of excited states based on TDDFT and to perform *ab initio* molecular dynamics. For this purpose, they have relied on a previously reported homology model³² and extracted the positions of the retinal atoms and its nearest neighbours. The focus of the present work is on the QM/MM nonadiabatic molecular dynamics based on a

multiconfigurational method. The purpose of this work is to do an initial step in closing the gap between the ultrafast spectroscopic studies and the computer simulations.

The paper is organized as follows: first we present the methodological details of the present work, then the computed excitation energies to validate our model. Starting from this optimized QM/MM model a relaxed scan and an excited state trajectory are reported and discussed in context of other retinal proteins and the ultrafast measurements.

Computational methodology

BPR is known for its high sequence identity with GPR,³³ which makes it the best suitable rhodopsin as a template. To generate a homology model of GPR we have taken the sequence of amino acids from a solution NMR study (PDB code: 4L6X).³⁴ Several templates were evaluated by creating homology models using the Swiss Model^{35,36} webserver and assessing the quality via MolProbity^{37,38} as well as visual inspection. Based on this assessment, chain B of the BPR crystal structure (PDB code 4JQ6)²⁷ was rated as the best template: It has a sequence identity of ca. 63 % and the resulting homology model did not show any Ramachandran outliers and a MolProbity score of 1.34. The latter is based on clashscore, rotamer, and Ramachandran evaluations and normalized to be on the same scale as X-ray resolution. Therefore, this homology model was used as the starting point in the present work.

In a first step, we added the retinal chromophore to the homology-derived model and linked it to the residue K231. In order to assign the protonation states of the titratable residues at pH 8, we employed the PDB2PQR web server.^{39,40} This pH value was used because it is close to the one of natural environment. The imidazole sidechain of H75 was reoriented to be inside the protein moiety according to the experimental study of Hempelmann *et al.*¹⁶ Further, 10 water molecules were added to the protein using the DOWSER program.⁴¹ The positions of all hydrogen atoms were optimized with the Amber16 program package.⁴²

The QM/MM optimization of the homology model was carried out using Amber16 interfaced to the Orca 4.0 program.^{43,44} The QM part, consisting of the retinal molecule and the lysine sidechain, was described with the BP86 functional^{45,46} in conjunction with the cc-pVDZ basis set⁴⁷. The hydrogen link atom was placed at the QM/MM boundary between C_α and C_β atoms of the K231 residue. The remaining protein was calculated with the Amber ff14SB force field,⁴⁸ while the water molecules were described using the TIP3P model.⁴⁹ All sidechains with at least one atom within 5 Å from the QM part were relaxed during the optimization.

The TD-BP86,^{45,46} TD-B3LYP,^{50,51} and TD-CAM-B3LYP⁵² vertical excitation energies were obtained using Orca 4.0⁴⁴ and the RI-CC2^{53,54} and RI-ADC(2)⁵⁵ calculations were realized with Turbomole 7.1.⁵⁶ All of the aforementioned calculations were carried out for the first five excited states with the cc-pVDZ basis set.⁴⁷ The excited state relaxed scan and nonadiabatic molecular dynamics simulations were performed for the first bright excited state using the Complete Active Space Self-

Consistent Field (CASSCF)⁵⁷ method as implemented in MOLCAS 8.1⁵⁸ developers version interfaced with Tinker 6.3.⁵⁹ All six bonding and six antibonding π -type orbitals were included into the active space. In order to reduce the computational cost of the more demanding CASSCF method the QM region was shortened. Therefore, the QM/MM boundary was placed between the C₆ and C₇ atoms of the Lys231 residue. The geometry was reoptimized using CASSCF on the ground state and keeping all MM atoms frozen at their positions. The relaxed scan on the excited state was computed by twisting the C₁₂-C₁₃-C₁₄-C₁₅ dihedral in steps of 5 degrees and relaxing the remaining degrees of freedom. The scan was computed for both senses of rotation of the C₁₃-C₁₄ bond. One excited state trajectory starting in the S₁ state without initial velocities, termed "FC trajectory" or "OK trajectory" in the literature,^{60–62} was propagated in time for 270 fs. The molecular dynamics was calculated using the Velocity-Verlet algorithm⁶³ and a time step of 1 fs. Surface hopping was realized via Tully's Fewest Switches scheme⁶⁴ using nonadiabatic coupling matrix elements based on wave function overlaps.⁶⁵ The integration of the electronic degrees of freedom was done employing 200 substeps for t and a decoherence correction of 0.1 Hartree was used.⁶⁶

Results and discussion

Geometry optimization and excitation energies

The homology model was subject to a QM(BP86/cc-pDVZ)/MM geometry optimization that included relaxation of the sidechains in the protein binding pocket. These sidechains were selected if at least one of its atoms is within 5 Å distance of the QM region. This procedure was chosen because the retinal geometry is refined in the crystallographic structure using force field parameters which are less accurate than the electronic structure methods. A similar approach was applied to other retinal proteins in QM/MM studies of the initial photochemistry^{67–69} and is justified by the fact that no conformational changes of the protein are expected on the time scale of the retinal photoisomerization. Indeed, a recent study on PR confirmed that first changes of the protein backbone appear after 500 fs.²²

In the optimized geometry of GPR the Schiff base group is oriented towards the D227, which is part of the counterion complex. The distance to D227 is with 2.8 Å (measured between iminium N and carboxyl C) significantly shorter compared to the 3.9 Å of the primary proton acceptor D97. This arrangement of the salt-bridge is adapted from the BPR crystal structure that was a template for the homology model. It was noted that BPR is missing a water molecule typically found in the counterion complex of bR (W402). Instead D227 forms a direct hydrogen bond to the iminium hydrogen. This attraction leads to an M helicity of the retinal backbone, characterized by -161° torsion of the N=C bond, slowly relaxing along the polyene chain with twisted C13=C14 (-169°) and C11=C12 (-178°).

Based on this model, we have calculated the vertical excitation energies using various electronic structure methods (Table 1, column 1). The protonation state of our model corresponds to

an alkaline environment, hence we take an experimental absorption maximum at 2.40 eV (516 nm) that was determined at pH 9.2.⁹ An experimental absorption maximum cannot be compared directly to a vertical excitation energy, hence we are not expecting to have a one to one correspondence with the computed values. Nevertheless, we observe that all methods are yielding excitation energies within a range of ± 0.25 eV. Even canonical TD-DFT methods like the GGA functional BP86 and the hybrid B3LYP yield results close to the measured value (2.26 eV and 2.42 eV, respectively), although both are notorious for their failure in describing charge transfer states. We have not found any dark states between the ground and the first excited state (Tables S1-S5 in ESI). The long-range corrected functional CAM-B3LYP produces a slightly higher excitation energy of 2.56 eV. These results are in agreement with a recent study by Buda *et al.*³¹ where B3LYP (2.43 eV) and CAM-B3LYP (2.72 eV) were employed as well. However, it should be noted that instead of a full QM/MM protein model a reduced cluster model was used in that study which contained rpSb with lysine, the counterion complex D97, D227, H75 and three water molecules which were treated at the DFT level and TZP basis set. Further, in that model the iminium moiety is hydrogen bonded to D97, instead of D227 as found in the BPR crystal structure. The comparison to the calculations in our study shows that this structural change has a small effect on the excitation energy.

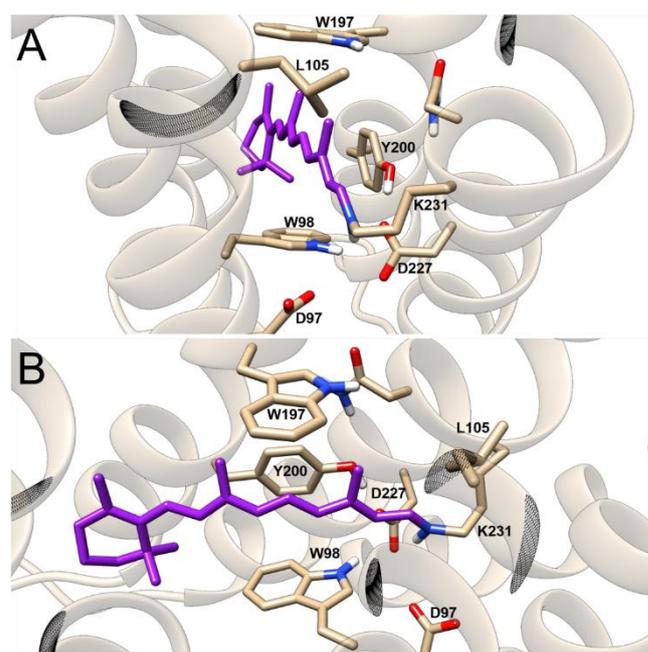


Figure 1. Optimized QM/MM model of GPR. The retinal depicted in purple is shown together with important amino acids of the binding pocket in licorice representation. A. View of along the Schiff base and the counter ion complex. B. View perpendicular to the retinal polyene chain.

In addition, we have tested RI-ACD(2) and RI-CC2 methods which have been previously applied for other retinal proteins.^{70,71} These methods result in 2.30 eV and 2.46 eV excitation energies, for RI-ACD(2) and RI-CC2, respectively. Further we have computed multiconfigurational and multireference energies. Surprisingly, SA2-CASSCF(12,12)

results in 2.46 eV which is close to the experimental value. This is a result of a fortuitous error compensation as the multireference perturbation theory to second order produces 2.65 eV, 2.27 eV and 2.66 eV for single state (SS), multistate (MS) and extended multistate (XMS) implementations of CASPT2. These methods should give more reliable values because they include dynamics electron correlation. The MS-variant is below the experimental value, while SS and XMS-CASPT2 are above. The values for the later should be more accurate because it is well known that MS leads to an artificially overestimated mixing of the electronic states at higher energy differences. MS-CASPT2 was originally designed to describe near-degeneracy situation and later replaced by XMS-CASPT2 that produces smoother surfaces close to degenerate energies.⁷² Although an extensive benchmark for XMS-CASPT2 vertical excitation energies is lacking, the agreement with SS-CASPT2 indicates less pronounced mixing. However, we should note that the difference is still within 0.25 eV to the experimentally determined 2.40 eV absorption maximum. An assessment of excitation energies from Thiel and co-workers has shown that CASPT2 energies are typically deviating by 0.2 eV from higher order methods.⁷³

Since we are relying on CASSCF for the calculation of the photoisomerization reaction, we have re-optimized the geometry of retinal using this method in the adapted QM/MM scheme as described in the method section. The resulting geometry has a significantly more pronounced bond length alternation pattern as documented in the literature.^{69,74–76} It means the double bonds are shorter while the single bonds are more extended compared to BP86. This leads to altered excitation energies (Table 1, column 2). The CASSCF-optimized geometry resulted in a large shift of 0.65 eV for TD-BP86, while the B3LYP and CAM-B3LYP results have changed by 0.10 and 0.15 eV, respectively. Surprisingly, RI-ADC(2) and RI-CC2 methods are less sensitive to the geometry and show a change below 0.1 eV. For CASSCF the excitation energies are shifted by almost 1 eV. The subsequent CASPT2 correction decreases the change in energy due to the different geometry but the difference is between 0.35–0.40 eV with respect to the first model. Hence, the structural changes have a notable effect on the excitation energies. However, in order to describe the photoisomerization that involves evolution on the ground and excited state, as well as passage close to a CI, we have to use a multiconfigurational method such as CASSCF.

Relaxed scan along the C₁₃=C₁₄ dihedral

The relaxed scan was generated by twisting the dihedral C₁₂-C₁₃=C₁₄-C₁₅ in steps of 5° in both senses of rotation starting from the FC point (Figure 2). The energy for each optimized geometry has been subsequently corrected by MS-CASPT2 calculations. The counter clockwise path leads to the energetically accessible S₀/S₁ conical intersection that is encountered shortly after -120° through a low barrier of 8 kcal/mol (MS-CASPT2). It should be noted that the transition state characterizing this barrier is energetically below the FC point, which means that a sufficient amount of kinetic energy should be available to overcome the barrier on the excited state. The

torsion approaching the S₀-S₁ degeneracy is smaller than the anticipated 90° from minimal conical intersection optimization. However, it is in line with previous computational studies that confirmed a 70° twist from planarity (starting from *cis* or *trans*). The nearly barrierless counter-clockwise path has been also discovered for other retinal proteins such as bR, Channelrhodopsin and *Anabaena* Sensory Rhodopsin.^{68,69,77–79} The structural changes associated with the excited state relaxation of the retinal are characterized by two dihedral angles with opposite senses of rotation. It can be seen in Figure 3 that the C₁₅=N bond is rotating in the opposite direction as compared to the C₁₃=C₁₄ bond, therefore reducing the space required to undergo a *trans-cis* isomerization. Such a mechanism was first proposed by Warshel as the so-called bicycle pedal isomerization.⁸⁰ However, similar to its refined “aborted bicycle pedal” mechanism formulated for bR, we also see only a high twist but not a full isomerization of C₁₅=N bond.⁸¹

The constrained excited state geometry optimization for torsions above -120° did not converge which is due the S₁ energy falling below the one of S₀. Hence, starting from -115° we have continued the optimization on the ground state (dotted line in Figure 2). The portion of the scan that follows the ground state after 90° shows the presence of a relatively high barrier due to steric repulsion of the hydrogen atom at C₁₄ and the hydroxyl hydrogen of the Y200 residue (gray area in Figure 2). However, towards the 13-*cis* isomer the retinal structure becomes planar again, the potential energy decreases and adapts a ground state configuration. This can also be seen from evolution of the bond length in this region (ESI). In comparison to the geometry at the FC point, a significant shortening of the single and elongation of the double bonds can be seen from the beginning of the scan due to the redistribution of the π-electron density in the excited state. However, after the S₁-S₀ transition and further optimization on the ground state the π-system recovers the alternation of the bond lengths, resembling the pattern of the ground state equilibrium geometry. The iminium hydrogen of the Schiff-base hydrogen bonded to the D227 residue throughout the scan and all retinal atoms, except for C₁₄ and C₁₅, remain close to their initial positions.

The clockwise path shows a decrease in the excited state energy in the first 10°, but reveals a significant potential barrier after further 30° that is ca. 10 kcal/mol (MS-CASPT2) higher than the FC point. Thus, the clockwise isomerization path is not accessible which is why we have decided to stop the scan in this direction.

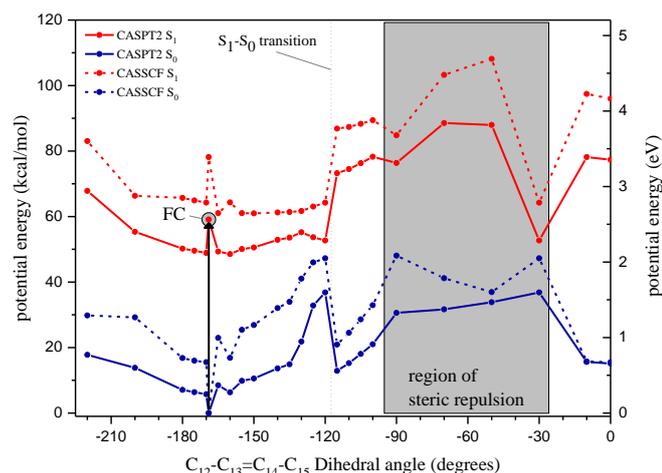


Figure 2. QM/MM relaxed scan along the C13=C14 dihedral of retinal in GPR. CASSCF (dashed line) and MS-CASPT2//CASSCF (solid lines) energies are shown along a C13=C14 torsion. The transition from S1 to S0 is indicated by a dotted line.

In recent studies of retinal proteins, the two senses of isomerization directions were attributed to the biexponential nature of the excited state decay.⁷⁷ The faster time constant was assigned to the path with a smaller energy barrier while the slower time constant was interpreted as the path with a higher barrier. However, although a biexponential decay has been also confirmed for GPR, our data indicates a unidirectional isomerization in GPR, similar to what has been described in the case of *Anabaena* Sensory Rhodopsin.⁷⁸

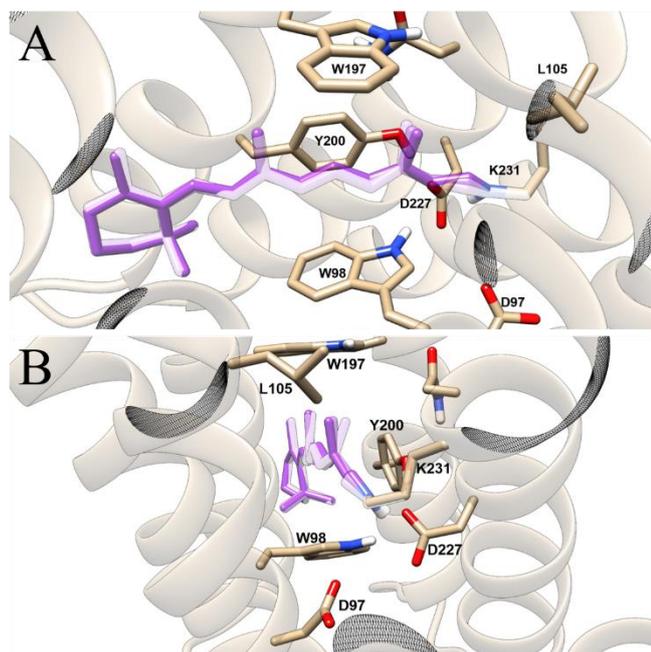


Figure 3. Geometrical changes along the relaxed scan. Three characteristic geometries are shown. The structures are decreasing in occupancy in the following order: Franck Condon point, 145° twist and 120° (close to the conical intersection). A. View of along the Schiff base and the counter ion complex. B. View perpendicular to the retinal polyene chain.

Excited State trajectory

In order to gain more insight into the excited state relaxation mechanism in GPR, we have carried out a nonadiabatic QM/MM molecular dynamics simulation. The trajectory was launched from the same FC point as in case of the relaxed scan (Figure 4). We have used zero initial velocities to run what has been referred to as a 0K-trajectory.^{60–62} A single trajectory is not sufficient to make statistics about lifetimes and the quantum yield of the photoisomerization, but it can provide some preliminary insight into the molecular mechanism. The evolution of the excited state trajectory is following a well known pattern.^{60,67,82–84} The first twenty femtoseconds of the MD can be characterized by an inversion of single and double bond length alternation at nearly planar geometry (Figure 4 B). Once the double bonds are becoming effectively like single bonds, an onset of the rotation about the double bonds is observed at around 30 fs of the dynamics (Figure 4 C). The largest twist is observed for the C12-C13=C14-C15 and C14-C15=N-C_e dihedral angles as already noticed in the relaxed scan. Although the evolution of the S1 potential energy is flat, the potential energy is slightly decreasing until the hop from S1 to S0 is observed at 226 fs. At the time of the S1-S0 transition, the value of the C13=C14 dihedral is -120°, in agreement with the distortion found in the relaxed scan. The C=N bond is twisting noticeably less, already rotating back to the *trans*-configuration at the time of the hop. In the following 30 fs it becomes apparent that both isomerizations are aborted returning to an all-*trans* retinal geometry. Inspecting the structural changes we have identified a collision between the hydrogen at C14 and the OH group of Y200 (Figure 5) that is responsible for an aborted isomerization. However, we cannot exclude the fact that our computational model with fixed amino acids in the binding pocket is responsible for this relaxation process on the ground state. However, we observe a space-saving rotation of the C13=C14 and C15=N double bonds of retinal in the excited state which leads to a conical intersection at 226 fs. This aborted bicycle pedal isomerization is consistent with the relaxed scan described above. The transition time of the 0K trajectory is on a similar time scale as the shortest of the two time constants for excited state

decay (200 fs) from the pump-probe experiments by Wachtveitl and co-workers²⁰.

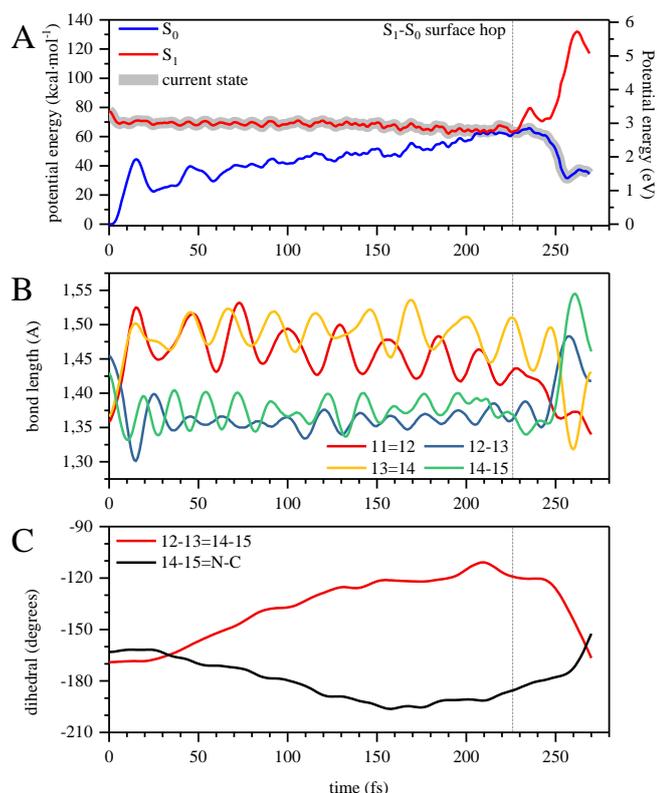


Figure 4. QM/MM nonadiabatic trajectory of GPR. Shown are the evolution of the energy (A); the bond lengths (B); and the dihedral angles (C).

Interestingly, Jung and coworkers⁸⁵ have recently characterized 18 variants of proteorhodopsin from the arctic ocean. They found 10 of them having an N200Y mutant and slower photocycle. However, this study has focussed on the late intermediates of the photocycle and didn't address the ultrafast isomerization.

Conclusions

We have presented a computational study of the initial photochemistry in GPR. For this purpose a homology model was constructed based on the available crystallographic structure of the related BPR. This model was refined with a QM/MM geometry optimization that provided a starting point for an excited state relaxed scan and a trajectory simulation. The relaxed scan along the $C_{13}=C_{14}$ located a conical intersection for a 70° distortion from planarity. Further, it revealed a detailed insight into the molecular level rearrangements: the $C_{13}=C_{14}$ rotation is accompanied by an anticlockwise twist of the $C_{15}=N$ double bond, also known as a bicycle pedal mechanism. This simultaneous motion in opposite directions minimizes the space required for the *trans-cis* isomerization in a tight protein pocket. Nevertheless, we found a high barrier while continuing the rotation on the ground state due to steric interaction between the hydrogen at C_{14} and the Y200 sidechain. This repulsion is enhanced in our model where the binding pocket is

held fixed. In a more flexible model such an interaction would probably lead to a delay of the isomerization without necessarily aborting it, which we plan to explore in a follow-up investigation.

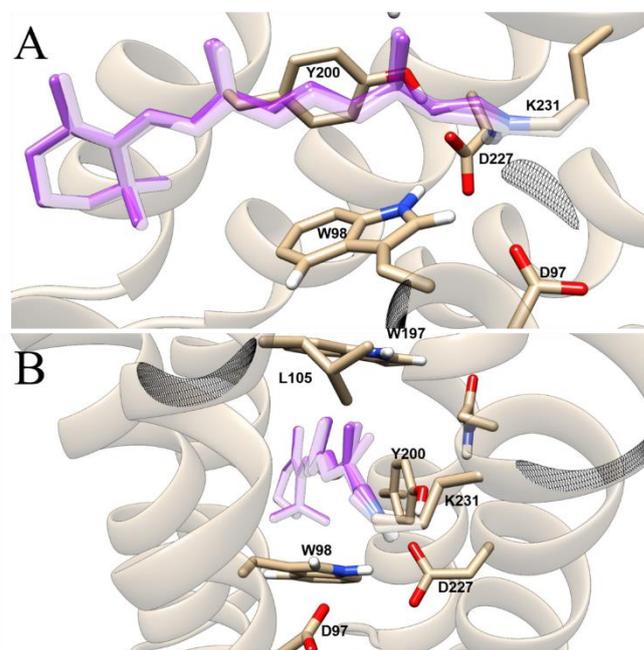


Figure 5. Geometrical changes along the QM/MM trajectory. Four characteristic geometries at different times are shown. The structures are decreasing in occupancy in the following order: Franck Condon point, 70 fs, 140 fs and 227 fs (transition to the ground state). A. View of along the Schiff base and the counter ion complex. B. View perpendicular to the retinal polyene chain.

The nonadiabatic QM/MM trajectory was computed starting from the same ground state geometry as the relaxed scan. The structural changes along the trajectory confirmed a bicycle pedal motion found in the relaxed scan computation. In addition, the trajectory revealed the timing of the events in the isomerization. The $C_{13}=C_{14}$ and $C_{15}=N$ double bond rotations start after the inversion of the bond length alternation in the first 20 fs. After 226 fs a transition to the ground state is observed, at this stage $C_{13}=C_{14}$ is twisted at -120° . Eventually, the repulsion between the C_{14} -H and hydroxyl group of Y200 prevents an isomerization to 13-cis-rpSb that leads to the starting geometry.

The presented work is the first computational study of the ultrafast photochemistry in GPR. It confirms the course of the isomerization reported for other, related rhodopsins. However, it gives insight into the specific molecular level events in GPR, like the increased repulsion between C_{14} hydrogen with the adjacent tyrosine residue. This understanding might lead to a design of specific Y200 mutants that could further delay or accelerate the reaction as well as impact the success rate of the isomerization. The PR variants with a N200Y mutation characterized by Jung and coworkers could be a natural example of such an adaptation.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 678169 "PhotoMutant"). Further, we would like to thank the Regional Computing Center of the University of Cologne (RRZK) for providing CPU time on the DFG-funded supercomputer CHEOPS and for their support. C.W. acknowledges support by the German Research Foundation (DFG) via a research scholarship (reference number: WI 4853/1-1).

References

- O. Béjà, L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich and E. F. DeLong, *Science*, 2000, **289**, 1902–1906.
- O. Béjà, E. N. Spudich, J. L. Spudich, M. Leclerc and E. F. DeLong, *Nature*, 2001, **411**, 786–789.
- C. Bamann, E. Bamberg, J. Wachtveitl and C. Glaubitz, *Biochim. Biophys. Acta - Bioenerg.*, 2014, **1837**, 614–625.
- O. P. Ernst, D. T. Lodowski, M. Elstner, P. Hegemann, L. S. Brown and H. Kandori, *Chem. Rev.*, 2014, **114**, 126–163.
- J. K. Lanyi, *Annu. Rev. Physiol.*, 2004, **66**, 665–688.
- T. Friedrich, S. Geibel, R. Kalmbach, I. Chizhov, K. Ataka, J. Heberle, M. Engelhard and E. Bamberg, *J. Mol. Biol.*, 2002, **321**, 821–838.
- R. A. Krebs, D. Dunmire, R. Partha and M. S. Braiman, *J. Phys. Chem. B*, 2003, **107**, 7877–7883.
- M. Lakatos, J. K. Lanyi, J. Szakács and G. Váró, *Biophys. J.*, 2003, **84**, 3252–3256.
- T. Köhler, I. Weber, C. Glaubitz and J. Wachtveitl, *Photochem. Photobiol.*, 2017, **93**, 762–771.
- B. Honig, U. Dinur, K. Nakanishi, V. Balogh-Nair, M. A. Gawinowicz, M. Arnaboldi and M. G. Motto, *J. Am. Chem. Soc.*, 1979, **101**, 7084–7086.
- F. Melaccio, N. Ferré and M. Olivucci, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12485.
- D. Man, W. Wang, G. Sabehi, L. Aravind, A. F. Post, R. Massana, E. N. Spudich, J. L. Spudich and O. Beja, *EMBO J.*, 2003, **22**, 1725–1731.
- M. T. Suzuki, O. Beja, L. T. Taylor and E. F. DeLong, *Environ. Microbiol.*, 2001, **3**, 323–331.
- M. O. Lenz, A. C. Woerner, C. Glaubitz and J. Wachtveitl, *Photochem. Photobiol.*, 2007, **83**, 226–231.
- M. K. Verhoefen, K. Neumann, I. Weber, C. Glaubitz and J. Wachtveitl, *Photochem. Photobiol.*, 2009, **85**, 540–546.
- F. Hempelmann, S. Hölper, M. K. Verhoefen, A. C. Woerner, T. Köhler, S. A. Fiedler, N. Pflieger, J. Wachtveitl and C. Glaubitz, *J. Am. Chem. Soc.*, 2011, **133**, 4645–4654.
- J. Herz, M. K. Verhoefen, I. Weber, C. Bamann, C. Glaubitz and J. Wachtveitl, *Biochemistry*, 2012, **51**, 5589–5600.
- J. Mao, N. N. Do, F. Scholz, L. Reggie, M. Mehler, A. Lakatos, Y. S. Ong, S. J. Ullrich, L. J. Brown, R. C. D. Brown, J. Becker-Baldus, J. Wachtveitl and C. Glaubitz, *J. Am. Chem. Soc.*, 2014, **136**, 17578–17590.
- C. E. Eckert, J. Kaur, C. Glaubitz and J. Wachtveitl, *J. Phys. Chem. Lett.*, 2017, **8**, 512–517.
- R. Huber, T. Köhler, M. O. Lenz, E. Bamberg, R. Kalmbach, M. Engelhard and J. Wachtveitl, *Biochemistry*, 2005, **44**, 1800–1806.
- M. O. Lenz, R. Huber, B. Schmidt, P. Gilch, R. Kalmbach, M. Engelhard and J. Wachtveitl, *Biophys. J.*, 2006, **91**, 255–262.
- J. J. Amsden, J. M. Kralj, L. R. Chieffo, X. Wang, S. Erramilli, E. N. Spudich, J. L. Spudich, L. D. Ziegler and K. J. Rothschild, *J. Phys. Chem. B*, 2007, **111**, 11824–11831.
- K. Neumann, M.-K. Verhoefen, I. Weber, C. Glaubitz and J. Wachtveitl, *Biophys. J.*, 2008, **94**, 4796–4807.
- A. Rupenyan, I. H. M. van Stokkum, J. C. Arents, R. van Grondelle, K. Hellingwerf and M. L. Groot, *Biophys. J.*, 2008, **94**, 4020–4030.
- A. Rupenyan, I. H. M. Van Stokkum, J. C. Arents, R. Van Grondelle, K. J. Hellingwerf and M. L. Groot, *J. Phys. Chem. B*, 2009, **113**, 16251–16256.
- E. Bühl, M. Braun, A. Lakatos, C. Glaubitz and J. Wachtveitl, *Biol. Chem.*, 2015, **396**, 1109–1115.
- T. Ran, G. Ozorowski, Y. Gao, O. A. Sineshchekov, W. Wang, J. L. Spudich and H. Luecke, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2013, **69**, 1965–1980.
- J. R. Hillebrecht, J. Galan, R. Rangarajan, L. Ramos, K. McCleary, D. E. Ward, J. A. Stuart and R. R. Birge, *Biochemistry*, 2006, **45**, 1579–1590.
- R. Rangarajan, J. F. Galan, G. Whited and R. R. Birge, *Biochemistry*, 2007, **46**, 12679–12686.
- J. Feng and B. Mertz, *Biochemistry*, 2015, **54**, 7132–7141.
- F. Buda, T. Keijer, S. Ganapathy and W. J. de Grip, *Photochem. Photobiol.*, DOI:10.1111/php.12800.
- S. Ganapathy, O. Bécheau, H. Venselaar, S. Frölich, J. B. van der Steen, Q. Chen, S. Radwan, J. Lugtenburg, K. J. Hellingwerf, H. J. M. de Groot and W. J. de Grip, *Biochem. J.*, 2015, **467**, 333–343.
- S. Y. Kim, S. A. Waschuk, L. S. Brown and K. H. Jung, *Biochim. Biophys. Acta - Bioenerg.*, 2008, **1777**, 504–513.
- S. Reckel, D. Gottstein, J. Stehle, F. Löhr, M. K. Verhoefen, M. Takeda, R. Silvers, M. Kainosho, C. Glaubitz, J. Wachtveitl, F. Bernhard, H. Schwalbe, P. Güntert and V. Dötsch, *Angew. Chemie - Int. Ed.*, 2011, **50**, 11942–11946.
- K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195–201.
- M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2014, **42**, 252–258.
- I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, J. Snoeyink, J. S. Richardson and D. C. Richardson, *Nucleic Acids Res.*, 2007, **35**, 375–383.
- V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and

- D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 12–21.
- 39 T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.*, 2004, **32**, W665–W667.
- 40 T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe and N. A. Baker, *Nucleic Acids Res.*, 2007, **35**, W522–W525.
- 41 L. Zhang and J. Hermans, *Proteins Struct. Funct. Genet.*, 1996, **24**, 433–438.
- 42 D. A. Case, R. M. Betz, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao and P. A. Kollman, AMBER 2016, University of California, San Francisco.
- 43 A. W. Götz, M. A. Clark and R. C. Walker, *J. Comput. Chem.*, 2014, **35**, 95–108.
- 44 F. Neese, *WIREs Comput. Mol. Sci.*, 2012, **2**, 73–78.
- 45 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
- 46 J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822–8824.
- 47 T. H. Dunning Jr, *J. Chem. Phys.*, 1989, **90**, 1007.
- 48 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
- 49 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 50 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 51 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
- 52 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
- 53 C. Hättig and F. Weigend, *J. Chem. Phys.*, 2000, **113**, 5154–5161.
- 54 C. Hättig and A. Köhn, *J. Chem. Phys.*, 2002, **117**, 6939–6951.
- 55 C. Hättig, *Adv. Quantum Chem.*, 2005, **50**, 37–60.
- 56 F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka and F. Weigend, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 91–100.
- 57 B. O. Roos, in *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry Part 2*, ed. K. P. Lawley, John Wiley & Sons, Inc., Hoboken, NJ, USA, 1987, vol. 69, pp. 399–445.
- 58 F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. Å. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata and R. Lindh, *J. Comput. Chem.*, 2016, **37**, 506–541.
- 59 J. W. Ponder and F. M. Richards, *J. Comput. Chem.*, 1987, **8**, 1016–1024.
- 60 L. M. Frutos, T. Andruniów, F. Santoro, N. Ferré and M. Olivucci, *Proc. Natl. Acad. Sci.*, 2007, **104**, 7764–7769.
- 61 C. García-Iriepa, M. Marazzi, F. Zapata, A. Valentini, D. Sampedro and L. M. Frutos, *J. Phys. Chem. Lett.*, 2013, **4**, 1389–1396.
- 62 M. Manathunga, X. Yang, H. L. Luk, S. Gozem, L. M. Frutos, A. Valentini, N. Ferré and M. Olivucci, *J. Chem. Theory Comput.*, 2016, **12**, 839–850.
- 63 W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *J. Chem. Phys.*, 1982, **76**, 637–649.
- 64 J. C. Tully, *J. Chem. Phys.*, 1990, **93**, 1061–1071.
- 65 S. Hammes-Schiffer and J. C. Tully, *J. Chem. Phys.*, 1994, **101**, 4657–4667.
- 66 G. Granucci and M. Persico, *J. Chem. Phys.*, 2007, **126**, 134114.
- 67 I. Schapiro, M. N. Ryazantsev, L. M. Frutos, N. Ferré, R. Lindh and M. Olivucci, *J. Am. Chem. Soc.*, 2011, **133**, 3354–3364.
- 68 I. Schapiro and S. Ruhman, *Biochim. Biophys. Acta - Bioenerg.*, 2014, **1837**, 589–597.
- 69 I. Dokukina and O. Weingart, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25142–25150.
- 70 K. Sneskov, J. M. H. Olsen, T. Schwabe, C. Hättig, O. Christiansen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7567–7576.
- 71 C.-M. Suomivuori, A. P. Gamiz-Hernandez, D. Sundholm and V. R. I. Kaila, *Proc. Natl. Acad. Sci.*, 2017, **114**, 7043–7048.
- 72 T. Shiozaki, C. Woywod and H.-J. Werner, *Phys. Chem. Chem. Phys.*, 2013, **15**, 262–269.
- 73 M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **128**, 134110.
- 74 F. Blomgren and S. Larsson, *J. Comput. Chem.*, 2005, **26**, 738–742.
- 75 M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim and M. Elstner, *J. Phys. Chem. B*, 2005, **109**, 3606–3615.
- 76 O. Valsson, C. Angeli and C. Filippi, *Phys. Chem. Chem. Phys.*, 2012, **14**, 11015–11020.
- 77 Y. Hontani, M. Marazzi, K. Stehfest, T. Mathes, I. H. M. van Stokkum, M. Elstner, P. Hegemann and J. T. M. Kennis, *Sci. Rep.*, 2017, **7**, 7217.
- 78 A. Strambi, B. Durbeej, N. Ferré, M. Olivucci and J. Michl, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 21322–21326.
- 79 P. Altoè, A. Cembran, M. Olivucci and M. Garavelli, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 20172–20177.
- 80 A. Warshel, *Nature*, 1976, **260**, 619–621.
- 81 A. Warshel and N. Barbooy, *J. Am. Chem. Soc.*, 1982, **104**, 1469–1476.
- 82 I. Schapiro, O. Weingart and V. Buss, *J. Am. Chem. Soc.*, 2009, **131**, 16–17.
- 83 D. Polli, P. Altoè, O. Weingart, K. M. Spillane, C. Manzoni, D. Brida, G. Tomasello, G. Orlandi, P. Kukura, R. A. Mathies, M. Garavelli and G. Cerullo, *Nature*, 2010, **467**, 440–443.
- 84 D. Polli, O. Weingart, D. Brida, E. Poli, M. Maiuri, K. M. Spillane, A. Bottoni, P. Kukura, R. A. Mathies, G. Cerullo and

M. Garavelli, *Angew. Chemie - Int. Ed.*, 2014, **53**, 2504–2507.

85 J. Y. Jung, A. R. Choi, Y. K. Lee, H. K. Lee and K. H. Jung, *FEBS Lett.*, 2008, **582**, 1679–1684.

Journal Name

ARTICLE

QM method	BP86		CASSCF	
Method	E _{vert} [eV] ([nm])	f	E _{vert} [eV] ([nm])	f
TD-BP86	2.26 (549)	1.41	2.91 (426)	0.94
TD-B3LYP	2.42 (512)	1.90	2.32 (534)	1.22
TD-CAM-B3LYP	2.56 (484)	2.15	2.70 (459)	1.72
CASSCF	2.46 (504)	1.90	3.36 (369)	1.01
SS-CASPT2	2.65 (468)	-	3.01 (412)	-
MS-CASPT2	2.27 (546)	2.23	2.64 (470)	1.22
XMS-CASPT2	2.66 (466)	2.31	3.09 (401)	1.37
RI-ADC(2)	2.30 (539)	1.81	2.37 (523)	1.26
RI-CC2	2.45 (506)	2.20	2.46 (504)	1.72

Table 1: Excitation energies in eV (nm) and oscillator strengths obtained from different quantum chemical methods for models optimized via DFT and CASSCF as the quantum chemical part in the QM/MM approach.. Experimentally, the first absorption maximum for GPR is found at ca. 516 nm at pH 9.2.⁹