Supplementary Material for PCCP

Product Formation in Rhodopsin by Fast Hydrogen Motions

Oliver Weingart*, Piero Altoè, Marco Stenta, Andrea Bottoni, Giorgio Orlandi and Marco Garavelli*

Supporting Information

S1. Computational details

The QM description of retinal in this study involves the six double bonds of the polyene chain and the CH_2-group of lysine at the CASSCF(10 π electrons/10 π orbitals)/6-31G* level of theory. The remaining retinal and protein atoms were treated with the Amber ff99 force field. [S1] The chromophore and two neighbouring water molecules were free to move in the simulations, while all other atoms were fixed at their positions in the protein crystal (PDBfile 1U19). [S2] Link-atoms were used to connect QM and MM regions. Starting from the optimized CASSCF/Amber ground state structure and a subsequent numerical frequency calculation, the retinal normal modes have been sampled at 300 K to obtain starting structures and velocities. [S3] High frequency stretching modes involving C-H and N-H were excluded from this procedure. QM/MM trajectories with surface hopping were calculated using the COBRAM interface combining Amber and Molpro. [S4] Details of the approach have been presented elsewhere [S5]. Other applications involving this methodology may be found in references [S6a-b].
For nuclear propagation the velocity Verlet integrator was used with a time step of 0.5 fs, reduced to 0.25 fs in the vicinity of an S_1/S_0 crossing. An equally weighed three root state average description was used throughout the entire calculation. For accurate calculation of state averaged (SA) CASSCF gradients Molpro's CPMCSCF routines were employed. Trajectories were brought to the ground state, when the scalar product of the S_1 and S_0 state coefficients indicated that a surface crossing had been passed. [S7] This method will underestimate the influence of surface hops with higher energy gaps, but allows for an accurate location of the crossing seam. The trajectories were propagated on S_0, until the final photoproduct could be identified from the evolution of the C11=C12 dihedral angles.

S2. Trajectory data and statistical analysis

After photoexcitation to S_1, the first motion is along stretching coordinates. C9=C10 and C11=C12 torsional coordinates are activated once the molecules are adapted to the excited state potential. At 70 fs into the simulation the majority of molecules has already a largely twisted C11=C12 bond (ca. 70° on average). Also the C9=C10 torsion has evolved by this time, but somewhat slower, from ca 170° to 210°. After ca. 100 fs the wavepacket branches into *cis*- and *trans* paths. In the *trans* branch the C9=C10 torsion persists at values around 220°, while the C11=C12 bond remains at ca. -140°.

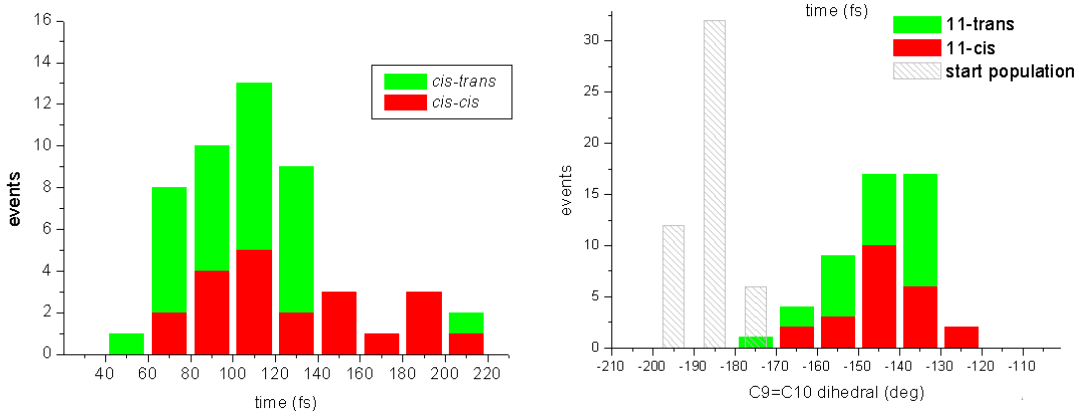

Figure S1. Statistical distribution of S_1 life times and C9=C10 torsion angles at decay point. The height of the red and the green bar graphs denote the number of trajectories reaching 11-*cis* and 11-trans photoproducts, respectively. The right graph includes the distribution or C9=C10 torsion at the start of the calculation.

Hops to S_0 occur preferably around ca. -90° C11=C12 torsion. Different from dynamics in isolated retinal models, where the seam persists all along the torsion coordinate, [S8, S9] the hopping angles in the protein are restricted to values between -70 and -110°. On average the C9=C10 angle twists by ca. 50° on S_1.

Figure S1 provides a statistical analysis of hopping times and C9=C10 torsion angles at hop, depicting additionally the partitioning of photoproducts. In total 29 out of the calculated 50 trajectories yield the all-*trans* photoproduct. Lifetimes cover a range from 40 to 200 fs. In the short living trajectories the probability of yielding the *trans* photoproduct is somewhat enhanced. A minor fraction of very fast trajectories (ca. 20%) arrives directly at a CI and decays to the ground state within 50-80 fs. The right graph in Figure S2 depicts the distribution of torsions at C9=C10 bond. Values at the beginning of the simulation are shown as shaded boxes. On average the C9=C10 angle twists by ca. 50° on S_1. The large motion and spread of this torsion angle is an indication of reduced steric strain and enhanced flexibility in this part of the chromophore.

The decision which photoproduct is finally generated will be taken in a short time frame after the surface hop, when the electronic situation of the molecule drastically changes. It takes some time before the molecules adapt to the new state surface and follow their topology. In this phase the formation of the π-bond can easily be influenced by the hydrogen atoms of the isomerising double bond, provided they have enough momentum and are able to change the preferred orientation of the π-orbital axis vector (POAV), which determines how the p-orbitals combine to form the π-bond. This can be achieved e.g. by inducing pyramidalisation at the corresponding carbon center. Although the POAV concept is static and geometry related, it is the dynamic change of the POAV in the “re-orientation”-phase of the molecule on the new surface, which finally influences product formation.

Figure S2 shows the analysis of motion at the hopping points with respect to the C10-C11=C12-C13 torsion and C11=C12 HOOP coordinates. The HOOP coordinate is calculated as the difference of H-

C=C-H and C-C=C-C torsions. A correlation between photoproduct and HOOP motion is visible, although smaller than in the corresponding H-C=C-H dihedral.

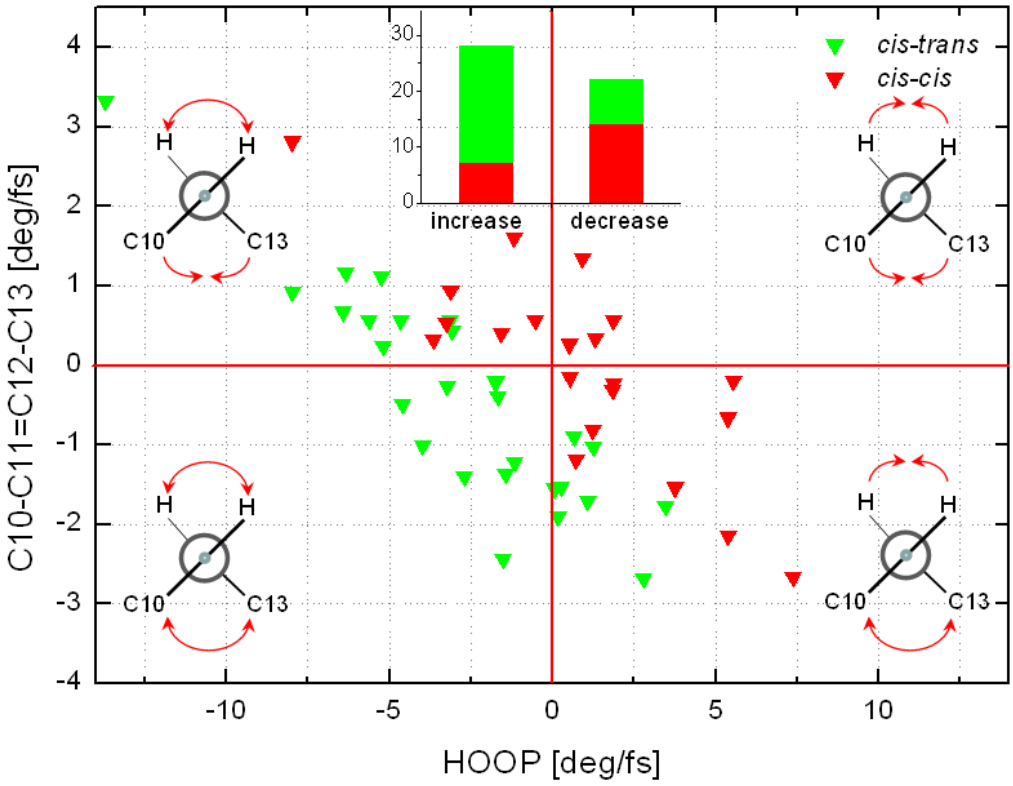

Figure S2. Gradient of the C10-C11=C12-C13 and C11=C12-HOOP torsion angles at the hopping points. Green squares (red triangles) denote trajectories yielding the trans photoproduct (cis educt). Negative (positive) values correspond to increasing (decreasing) norm angles.

The derivative coupling vector at the point of surface crossing gives information on the motion, which most effectively leads the molecule to and from the intersection. In several trajectories we find participation of hydrogen out-of-plane motion, leading to strong pyramidalisation at the C11 and C12 carbon centers. Figure S3 shows this vector for three selected hopping structures in massweighted units.

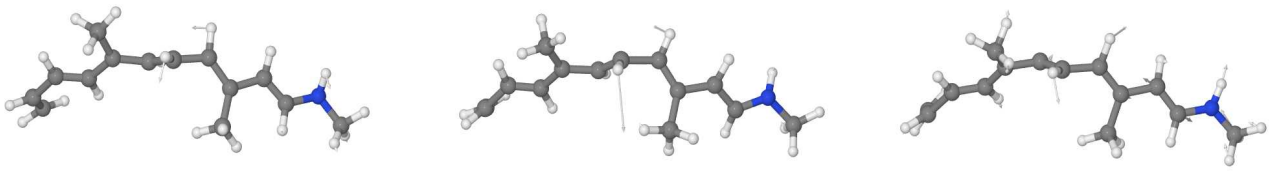

Figure S3. Derivative coupling vector at surface hopping point in selected trajectories. Vector shown in massweighted units.

S.3. Energy storage

In the following we analyse the energy of the intermediate photoproduct in order to get an estimate on the energy stored in this structure. After relaxation of the hot ground state all-*trans* photoproduct, a large

amount of the kinetic energy collected in the excited state is stored in molecular deformations of the chromophore, especially by twists of the double bonds. Experimental values for this energy storage in the first stable photoproduct bathorhodopsin (that is reached in a ps timescale) are 30-40 kcal/mol [S10]. The energy is released in later steps and used to move the helices, which results in activation of the protein for its coupling partner transducin. When comparing computations for Rh and geometry-optimised batho-structures this energy difference is usually underestimated [S11]. Dynamical effects may apply, which render impossible a proper treatment of energy storage with static optimisation methods. Our trajectory set does not cover the ps timescale needed to reach bathorhodopsin. However, it is possible to deliver an estimate of the energy stored in the primary photo-rhodopsin intermediate that is indeed reached within 200 fs by many (ca. 60%) of our trajectories: averaging over those trajectories reaching the all-*trans* photoproduct from 50 fs after the surface hop until the end of the calculations yields an approximate estimate of the energy stored in the photorhodopsin structure. Average energies were calculated for 5085 data points that fulfil the conditions mentioned above. The difference to the average ground state energy of the 11-*cis* starting configuration is ca. 46,5 kcal/mol. As expected, and remarkably, this value is significantly higher than that reported for bathorhodopsin. An average over longer (1 ps at least) trajectories would be necessary to get reliable data on the energy stored in bathorhodopsin.

S.3. Limitations

Due to the models and methods used, some limitations apply for the calculations performed in this work. In the following we will discuss some of these limitations and their influence on the calculated data.

The CASSCF method used in the calculation of the trajectory set forms a compromise between accuracy and computational effort. Application of the more reliable, dynamically correlated CASPT2 method is currently only possible for very small systems [S12], but not for the complex rhodopsin protein. Figure S4 shows a comparison of the S_1 torsional profile calculated at the 3 root CAS(10/10)/6-31G* level (i.e., the same level employed for the dynamics) with the corresponding re-evalutaed CASPT2 profile. The CASPT2 surface is less steep in the region between the Franck Condon point (FC) and the S_1 minimum, i.e. CASSCF will overestimate the reaction rate at the start of the photoreaction. Both energy profiles are barrierless, and the CASSCF conical intersection point is almost reproduced in CASPT2. The extrapolated CI point at this level appears at a slightly lower torsion angle. The part from the S_1 minimum to ca 70° is almost parallel in steepness for both methods, thus the CASSCF dynamics will qualitatively reproduce the excited state features of the dynamically correlated surface.

The fixed crystal structure of the protein in our simulation is considered as a mean representation of the experimental positions of the atoms, although at low temperature (100 Kelvin). This choice is justified by the extremely short timescale of the retinal photoreaction, which prevents thermalization of the protein. This approximation is further supported by results of the Schulten group which showed, that the excited state dynamics of retinal are not affected by protein mobility [S13]. It was also considered for practical reasons: we apply an electrostatic embedding scheme, i.e the chromophore is polarised by the MM environment. The program we have used has certain limitations on the number of electrostatic field derivatives that can be processed simultaneously. Furthermore, a mobile protein environment may result in a wrong ordering of S1 and S2 energy states at the CASSCF level, leading to unrealistic simulations. Protein motion and relaxation will of course become important in the next and longer steps of the reaction, namely the formation of bathorhodopsin and the meta derivatives, where part of the kinetic energy accumulated in the excited state is transferred to the protein.

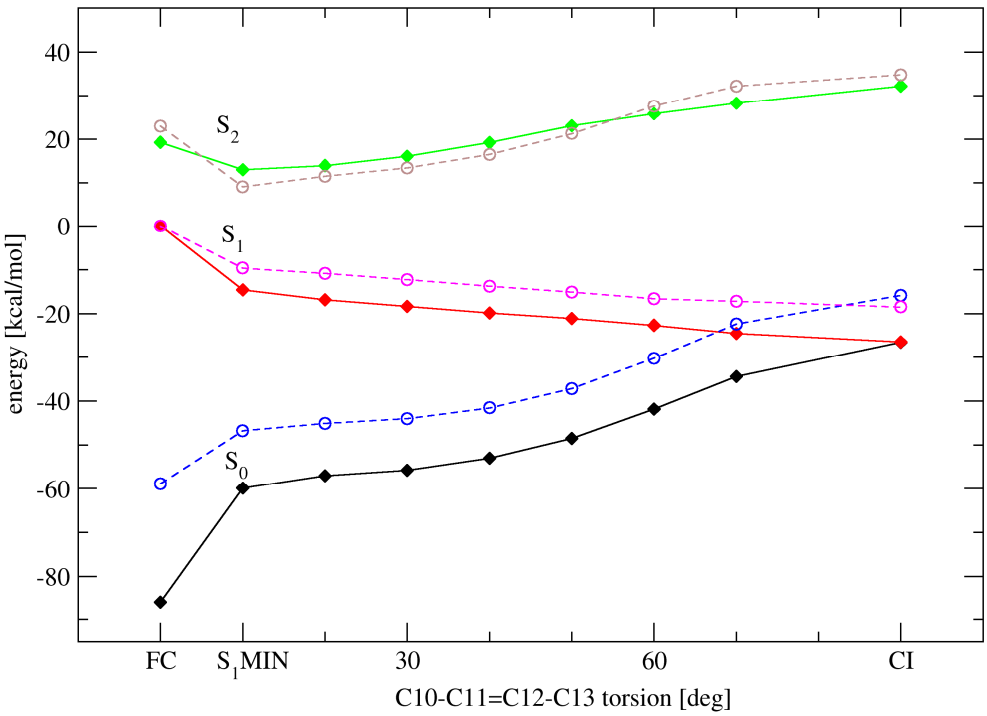

Figure S4: CASSCF(10/10)/6-31G* calculated torsion profile (full lines) from FC-point to the CI (90°) re-evaluated at the CASPT2 level (broken lines)

Following the argumentation for product formation, the surface hopping method applied is responsible for the reduced quantum yield and a longer excited state life time (110 fs vs. 80 fs of the experiment). The use of this method however pointed out the importance of the conical intersection seam in the photoreaction of retinal in opsin. The seam is unavoidable for the excited molecules: all of the trajectories finally hop to the ground state at a surface crossing.

References

[S1] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, *J. Comp. Chem.* **2005**, 26, 1668

[S2] T. Okada, M. Sugihara, A.N. Bondar, M. Elstner, P. Entel, V. Buss, *J. Mol. Biol.* **2004**, 342, 571-583

[S3] P. Altoe, M. Stenta, A. Bottoni, M. Garavelli, *Theoret. Chem. Acc.* **2007**, 118, 219

[S4] MOLPRO, version 2008.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz et al.

[S5] C. S. Sloane, W. L. Hase, *J. Chem. Phys.* **1977**, 66, 1523-1533.

[S6] a) M. Stenta et al. *J. Phys. Chem. B* **2008**, 112, 1057-1059. b) G. Tomasello et al. *J. Am. Chem. Soc.* **2009**, 131, 5172-5186

[S7] O. Weingart, I. Schapiro, I., V. Buss, *J. Phys. Chem. B* **2007**, 111, 3782-3788

[S8] O. Weingart, A. Migani , M. Olivucci , M. A. Robb, V. Buss, P. Hunt *J. Chem. Phys.* **2004**, 108, 4685-4693

[S9] A. Migani, M. A. Robb, M. Olivucci, *J. Am. Chem. Soc.* **2003**, 125, 2804-2807

[S10] A. Cooper, *Nature* **1979**, 282, 531–533

[S11] M. Schreiber, M. Sugihara, T. Okada, V. Buss, *Angewandte* **2006**, 45, 4274-4277

[S12] H. Tao, B.G. Levine, T.J. Martinez, *J. Phys. Chem. A* **2009**, 113, 13656–13662

[S13] S. Hayashi, E. Taikhorshid, K. Schulten, *Biophys. J.* **2009**, 96, 403-416