

# PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

## ARTICLE

# Conversion of Light-Energy Into Molecular Strain in the Photocycle of the Photoactive Yellow Protein

Ana P. Gamiz-Hernandez, Ville R. I. Kaila

X Cite this: DOI:  
10.1039/x0xx00000x

Received 00th January 2015,  
Accepted 00th January 2015

DOI: 10.1039/x0xx00000x

www.rsc.org/

The Photoactive Yellow Protein (PYP) is a light-driven photoreceptor, responsible for the phototaxis of halophilic bacteria. Recently, a new short-lived intermediate ( $pR_0$ ) was characterized in the PYP photocycle using combined time-resolved X-ray crystallography and density functional theory calculations. The  $pR_0$  species was identified as a highly contorted *cis*-intermediate, which is stabilized by hydrogen bonds with protein residues. Here we show by hybrid quantum mechanics/classical mechanics (QM/MM) molecular dynamics simulations, and first-principles calculations of optical properties, that the optical shifts in the early steps of the PYP photocycle originate from the conversion of light energy into molecular strain, stored in the  $pR_0$  state, and its relaxation in subsequent reaction steps. Our calculations quantitatively reproduce experimental data, which enables us to identify molecular origins of the optical shifts. Our combined approach suggests that the short-lived  $pR_0$  intermediate stores  $\sim 1/3$  of the photon energy as molecular strain, thus providing the thermodynamic driving force for later conformational changes in the protein.

## 1. Introduction

The photoactive yellow protein (PYP) is a small (14-kDa) water-soluble protein found in purple photosynthetic bacteria that live in halophilic environments.<sup>1</sup> PYP functions as a blue-light receptor, in which the light-absorption triggers a signaling cascade that switches the rotational direction of the bacterial flagellum, and directs the bacteria to swim away from potential harmful high energy light.<sup>2, 3</sup> PYP has become a key photobiological model system for studying light-absorption induced changes in proteins, due to its small size but relatively complex photocycle.<sup>4</sup> The photoactive center of PYP comprises a *p*-coumaric acid (*p*CA) chromophore that is covalently linked to a cysteine residue (Cys-69) via a thioester bond (Figure 1). The *p*CA is stabilized by hydrogen bonds from Tyr-42, Glu-46 and Thr-50, whereas the backbone of Cys-69 stabilizes the chromophore in early intermediates of the photocycle. Arg-52 shields the chromophore from the surroundings, and may electrostatically stabilize the anionic chromophore. In the PYP dark state (*p*G), shown in Figure 1, the C2=C3 double bond of *p*CA is in a *trans*-conformation, whereas light-absorption triggers a *trans*-to-*cis* isomerization that leads to subsequent formation of redshifted (*p*R) and blueshifted (*p*B) photocycle intermediates on picosecond – millisecond time scales.<sup>5-9</sup> Recently, Schotte *et al.*<sup>10</sup> characterized a new photocycle intermediate using picosecond time-resolved X-ray crystallography and density functional theory (DFT) calculations.<sup>10, 11</sup> This new short-lived state,  $pR_0$ , appeared on a 150 ps timescale (Figure 2B), and was characterized as a highly strained *cis*-species with the double bond *ca.* 30° out of plane. The  $pR_0$  state was found to further relax into  $pR_1$ ,  $pR_2$  and *p*B states,<sup>10, 12</sup> that were found to be structurally and kinetically

consistent with previously refined structural intermediates obtained from time-resolved X-ray crystallography.<sup>12-15</sup> Recently Jung *et al.*<sup>16</sup> also characterized a new short-lived species,  $I_T$ , forming on time-scales similar to the  $pR_0$  state. In contrast to Schotte *et al.*,<sup>10</sup> however, the refined species was suggested to be a transition state structure between *p*G and later intermediates, with the C2=C3 double bond refined to *ca.* 90° out of plane. Jung *et al.* also used DFT calculations to study the likelihood of their species, but in contrast to their refined X-ray structure, the obtained DFT-structure was in close structural agreement with the  $pR_0$  species described by Schotte *et al.*<sup>10</sup> Differences between these two structures were recently discussed Kaila *et al.*<sup>11</sup> and Jung *et al.*<sup>17</sup> However, the focus of this study concerns optical tuning effects in the PYP photocycle studied by QM/MM calculations.

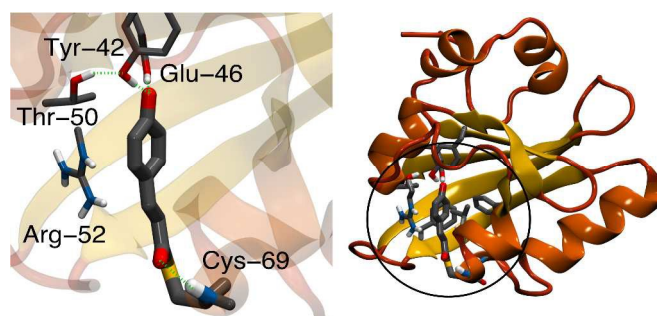
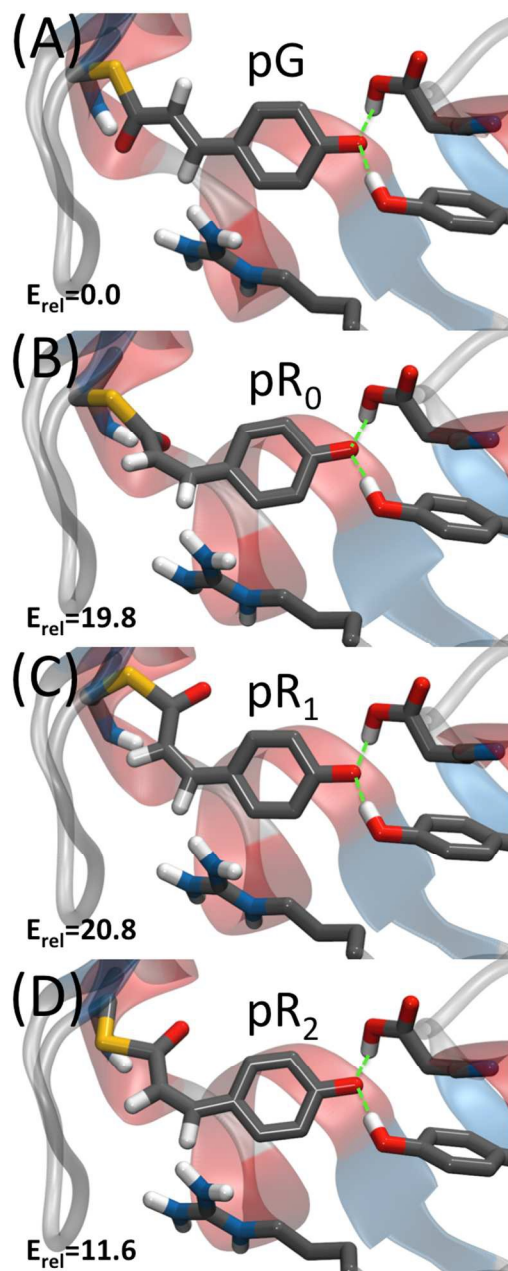


Figure 1. The structure of the *p*CA chromophore in the PYP dark state (*p*G). Left: *p*CA and its hydrogen bonding partners. Right: The structure of PYP showing the chromophore-binding region, modeled at the QM level of theory.

Proteins can tune the color of their biochromophores by electrostatic polarization and/or by imposing strain effects on the chromophore.<sup>18-23</sup> By stabilizing the electronic ground state or excited state, the absorption can be shifted towards the blue (higher energy) or red (lower energy) regime of the spectrum. The strain energy stored in the recently discovered pR<sub>0</sub> species is expected to lead to a destabilization of the electronic ground state, which in turn would further redshift the optical properties. Moreover, relaxation of the strain in the subsequent photocycle steps would be expected to blueshift the absorption.



**Figure 2.** Structure of photocycle intermediates. The relative energy ( $E_{\text{rel}}$  in  $\text{kcal}\cdot\text{mol}^{-1}$ ) stored within the pCA chromophore is also shown. The  $\text{C}_2=\text{C}_3$  double bond in pCA is *trans*-in the pG state and *cis* in all pR states. This double bond is strained in pR<sub>0</sub>, which upon breaking the hydrogen bond to Cys-69 converts to pR<sub>1</sub>, which is much more planar and has less strained. The transition from pR<sub>1</sub> to pR<sub>2</sub> involves a rotation of the  $\text{C}\beta\text{-S}\gamma\text{-C}_1\text{-C}_2$  dihedral angle, which transforms from *syn-syn* to *anti-anti*, and further reduces strain in the pCA.

Computation of optical properties of photobiological systems from first-principles is extremely challenging due to the complexity of the biological “active site” region, which normally limits the application of accurate correlated *ab initio* methods.<sup>19, 20, 24-28</sup> The effect of the protein environments can be accounted for by employing *e.g.* hybrid quantum mechanical/molecular mechanics (QM/MM)<sup>19, 20, 24, 27-31</sup> or frozen density embedding approaches.<sup>22, 32, 33</sup> Previous studies also suggest that nearby environmental effects must explicitly be considered at a quantum chemical level of theory<sup>34, 35</sup> or by accurate polarizable models<sup>36</sup> to obtain quantitatively accurate results.

To study the structure and dynamics of the early photocycle intermediates in PYP, we use here a combination of QM/MM molecular dynamics simulations, first-principles computation of optical properties at second-order approximate coupled cluster (CC2) and time-dependent density functional (TDDFT) theory levels to elucidate how the photon energy is converted into molecular strain in the PYP photocycle, and to characterize the dynamics of the subsequent intermediates.

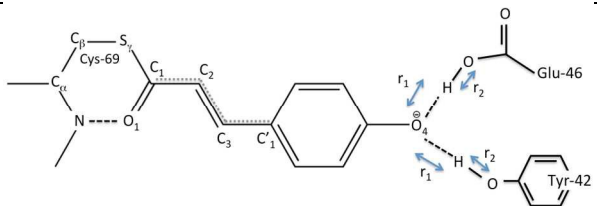
## 2. Computational Methods

Time-resolved X-ray structures of the pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> states (PDB ID: 2ZOH, 4B90, 4BBT and 4BBU) were obtained from the Brookhaven Protein Databank and used as starting points for the QM/MM simulations. We also performed QM/MM simulations on the refined I<sub>T</sub> structure by Jung *et al.*<sup>16</sup> (PDB ID: 4I38) to study the X-ray model dependence of our data. The QM region (143 atoms) included the pCA chromophore, Tyr-42, Glu-46, Thr-50, Arg-52, Phe-62, Val-66, Ala-67, Pro-68, Cys-69, Thr-70 and Phe-96, and was described at the D-BP86/def2-SVP level of theory. The MM region, comprising the surrounding protein residues and water molecules, was described using the CHARMM27 force field.<sup>37</sup> Each state was structure optimized, followed by 5 ps QM/MM MD simulations using a time step of 1 fs at 310 K. From the QM/MM trajectories, we extracted 100-500 snapshots for each state along the MD trajectories for calculation of ground state and optical properties. The absorption spectrum of the early PYP photocycle states was computed at the time-dependent density functional theory (TDDFT) level using CAM-B3LYP/def2-SVP, and at the second-order approximate coupled cluster level (CC2/def2-TZVP). The first six vertical excitation energies (VEEs) at TDDFT level of theory were computed for each snapshot leading to evaluation of 5 000 VEEs per method (10 000 VEEs in total), whereas at the CC2 level, we computed the first two VEEs. To speed up the evaluation of the VEEs at the CC2 level, we employed the reduced virtual space (RVS) approach<sup>38</sup> where virtual orbitals above 70 eV from the highest-occupied molecular orbital (HOMO) were disregarded from in the correlation calculations, leading to a RVS<sup>38</sup> cutoff error of 0.02 eV relative to frozen core CC2 calculations (Figure S1). For the isolated chromophore, the CC2 and CAM-B3LYP produced similar VEE profiles, but CAM-B3LYP was found to systematically blueshift the absolute excitation energies relative to CC2 (Figure S2). All CAM-B3LYP profiles were therefore redshifted by a 0.37 eV offset-factor obtained from the CC2 fit. For calculation of optical properties, we reduced the size of the QM region to explicitly include the pCA chromophore, Cys-69, Tyr-42, Glu-46 and Arg-52, for the TDDFT calculations, and the pCA chromophore for CC2 calculations, while the remaining protein and water atoms were described using point



charges obtained from the QM/MM trajectories. For the estimation of electrostatic-induced shifts, we calculated the optical spectra of the chromophore along the QM/MM trajectories, by removing all protein residues around the chromophore but by keeping the chromophore structure unchanged. Strain-induced optical shifts were calculated as the difference between the absorption maximum of the isolated (*trans* or *cis*) chromophore structures obtained from gas-phase and the protein simulations, without considering the electrostatic polarization of the protein environment. The absorption spectra were calculated as an average of 500 VEEs obtained from the MD trajectories and weighted with their respective oscillatory strengths for each VEE calculation. The distribution of weighted VEE along the trajectories was fitted with a Gaussian width of  $100 hc / E_h$  eV, where  $E_h$  is the excitation energy in eV, and combined to give a spectra by scaling the maximum intensity to 1 or by matching to the maximum experimental intensities. The two first excited states were included in the calculations due to larger oscillator strengths at TDDFT level. The spectra of each state were also calculated by computing histograms of the observed VEEs along the MD trajectories, and by normalizing the probability distribution. We obtain very similar spectra using both the Gaussian fitting and the histogram approaches, showing that the calculated line shapes arise from the statistical sampling of the ground state dynamics. Both approaches were found to give similar absorption spectra. QM/MM simulations were performed using the Q-Chem/CHARMM,<sup>37, 39</sup> TDDFT and CC2 (RVS) calculations were performed with Q-Chem v4.1<sup>40</sup> and TURBOMOLE v6.3,<sup>41</sup> respectively.

Table 1. Comparison of geometrical parameters between crystal structure and average ( $\pm$  standard deviation) values obtained from 5 ps QM/MM trajectories of the four photocycle states.



	Dihedral ( $^\circ$ )			Hydrogen bond ( $\text{\AA}$ )			
	$C_1C_2C_3C_1'$	$O_1C_1C_2C_3$	$S_1C_1C_2C_3$	Y42-O4'	E46-O4'	T50-Y42	C69-O1
pG <sup>C</sup>	168	-10	176	2.52	2.56	2.91	2.79
pG <sup>Q</sup>	165 $\pm 8$	0 $\pm 10$	182 $\pm 11$	2.52 $\pm 0.08$	2.64 $\pm 0.10$	2.73 $\pm 0.10$	3.00 $\pm 0.26$
pR <sub>0</sub> <sup>C</sup>	33	29	-154	2.59	2.73	2.95	3.05
pR <sub>0</sub> <sup>Q</sup>	28 $\pm 18$	29 $\pm 12$	-153 $\pm 10$	2.55 $\pm 0.09$	2.69 $\pm 0.11$	2.82 $\pm 0.15$	2.85 $\pm 0.15$
pR <sub>1</sub> <sup>C</sup>	-8	4	-174	2.57	2.64	2.92	4.91 <sup>a</sup>
pR <sub>1</sub> <sup>Q</sup>	11 $\pm 13$	15 $\pm 10$	-159 $\pm 9$	2.53 $\pm 0.08$	2.68 $\pm 0.13$	2.73 $\pm 0.09$	4.10 $\pm 0.26$
pR <sub>2</sub> <sup>C</sup>	2	4	-175	2.47	2.61	2.92	3.88
pR <sub>2</sub> <sup>Q</sup>	16 $\pm 11$	5 $\pm 11$	-178 $\pm 9$	2.54 $\pm 0.08$	2.69 $\pm 0.12$	2.78 $\pm 0.11$	3.49 $\pm 0.32$

<sup>a</sup> Hydrogen bond broken.

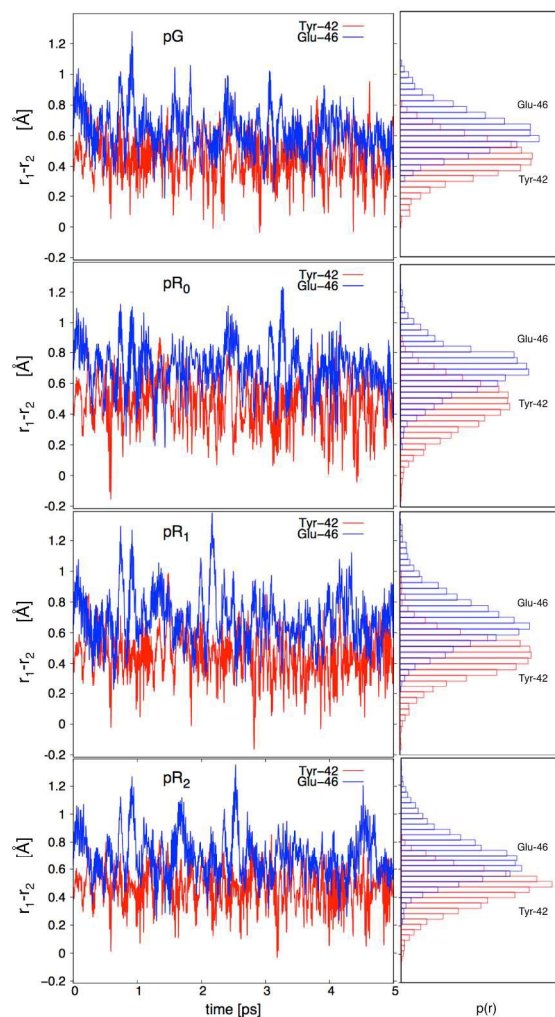


Figure 3. Dynamics of hydrogen bonding distances in the pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> states. The proton coordinate  $r = r_1$  (TyrO-H/GluO-H) -  $r_2$  (pCA:O $\cdots$ H) (in  $\text{\AA}$ ), with the Tyr-42 and Glu-46 bonds shown in red and blue, respectively.  $r < 0$  indicates that the proton is closer to pCA than to Glu-42/Tyr-46. Right inset: probability distributions of  $r$ , showing that the Tyr-42 is shorter relative to the Glu-42 bond.

### 3. Results and Discussion

#### Ground State Dynamics of PYP Photocycle Intermediates

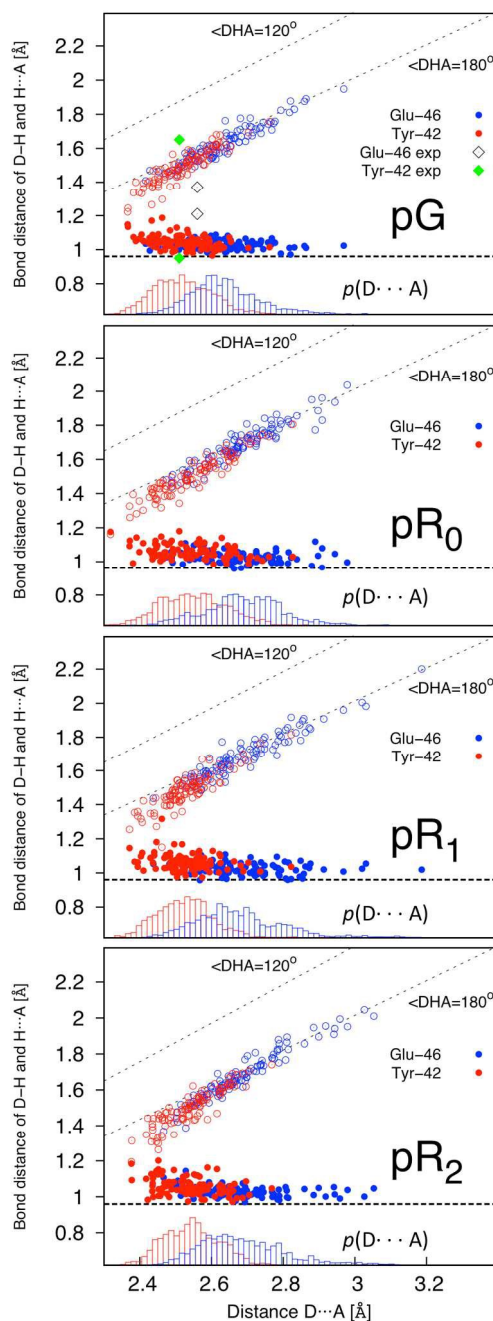
Our QM/MM molecular dynamics simulations suggest that all studied photo-cycle intermediates; pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> remain stable on picosecond time-scales. Figure 2 shows how the energy stored in the dihedral bonds of the chromophore is released during the pR<sub>0</sub>  $\rightarrow$  pR<sub>2</sub> transition, in which the dihedral angle around the C<sub>2</sub>=C<sub>3</sub> double bond relaxes from a  $\sim 30^\circ$  strained *cis*-conformation in pR<sub>0</sub> to planar species in pR<sub>1</sub> and pR<sub>2</sub>. The pR<sub>0</sub> state stores *ca.* 20 kcal mol<sup>-1</sup> of strain energy relative to the pG state, which corresponds to about 1/3 of the photon energy (446 nm = 64 kcal mol<sup>-1</sup>). The pR<sub>0</sub> state is stabilized by a hydrogen bond between the carbonyl group of

pCA and the backbone of Cys-69 (Figure 2), as suggested by Schotte *et al.*<sup>10</sup> Consistent with time-resolved X-ray crystallographic experiments, we also observe a transient formation of a planar pR<sub>1</sub>-like species during the trajectory, suggesting that the pR<sub>0</sub>/pR<sub>1</sub> species may reside in (quasi)-equilibrium during turnover (see Movie S1, C-D)). Although the strain energy of the pR<sub>1</sub> state is lower relative to pR<sub>0</sub>, the intrinsic energy of the chromophores in the two states are similar, due to dissociation of the backbone hydrogen bond between pCA and Cys-69 (Figure 2). In the pR<sub>2</sub> state, we observe *ca.* 5-10 kcal·mol<sup>-1</sup> release of strain energy relative pR<sub>1</sub> by further relaxing the C<sub>β</sub>-S<sub>γ</sub>-C<sub>1</sub>-C<sub>2</sub> dihedral angle (Table 1).

We also performed QM/MM MD simulations on the I<sub>T</sub> structure of Jung *et al.*<sup>16</sup> to study how the dynamics depends on the employed X-ray model. Consistent with our earlier observations<sup>11</sup> we find that the refined I<sub>T</sub> structure is unstable and the dihedral angle around the C<sub>2</sub>=C<sub>3</sub> rapidly relaxed from 80° to 34°, and stays in this conformation during the 5 ps QM/MM simulation trajectory (Figure S4). The two X-ray models thus result in very similar conformational dynamics. The hydrogen bonding distances between pCA, and Tyr-42 and Glu-46, are shown in Figure 3 for the four studied photocycle intermediates. We observe short hydrogen-bonding distances (2.5 Å, Table 1) and partial proton transfer between Tyr-46 and the pCA, which becomes more pronounced in the pR<sub>1</sub> and pR<sub>2</sub> states as strain energy is released from the chromophore. A possible reason is that the proton affinity of the pCA increases with the decrease in strain due to flow of electron density from the double bond to the phenol ring (Figure S3). The hydrogen bond between Glu-46 and the pCA is also short (2.6 Å), but in contrast to Tyr-42, we do not observe proton transfer between the two former residues.

Yamaguchi *et al.*<sup>42</sup> refined the position of hydrogen (deuterium) atoms in the pG state using neutron diffraction experiments of deuterated samples, and observed an unusually short hydrogen (deuterium) bond of 1.38 Å between the Glu-42 and the pCA. Saito and Ishikita<sup>43, 44</sup> studied the energetics of the proton transfer between Glu-42 and pCA by hybrid QM/MM optimizations, and observed short distances between the residues only upon dissociation of Tyr-42 from pCA. They also found a slightly uphill proton transfer barrier between Tyr-42 and the pCA, which seems consistent with our dynamics simulations, where the proton transiently fluctuates towards the pCA (Figure 3). To study the distance distribution of hydrogen bonds observed in the different photocycle states, we computed the correlation between the proton donor (*D* - phenolic/carboxylic oxygen of Tyr or Glu) and proton acceptor (*A* - phenolic oxygen of pCA) distances, with the *A-H* (acceptor-proton) and *D-H* (donor-proton) distances. This correlation plot, shown in Figure 4, suggests that short hydrogen bonding distances, indicated by blue dotted lines in Figure 4, are observed for Tyr-42, but not for Glu-46. Our simulations further suggest the short distances are observed both for the *D-H* (closed circles) and *A-H* (open circles) bonds when the *A-D* distance is < 2.5 Å. It is, however, possible that the experimentally observed short distances between the pCA and Glu-46 could arise from dynamically averaged fluctuations of multiple positions relative to the static heavy atom positions obtained from the X-ray structure in the neutron diffraction

refinement. Such short bonds have, *e.g.*, been observed in high resolution X-ray structures of aquaporin.<sup>45</sup>



**Figure 4.** Correlation of the *D-H* and *A-H* bond with the *D-A* distance (in Å) in the pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> states. The *D*, *H*, and *A*, refer to the phenolic/carboxylic oxygen of Tyr-42/Glu-46; the proton between these residues and the pCA; and the phenolic proton of the pCA, respectively. Distances obtained for the Glu and Tyr bonds are shown in blue and red, respectively, with bonds originating from the *D-H* distance (in open circles) and *A-H* distance (in closed circles). Lower inset in each subfigure shows probability distribution of the *D-A* distance. The hydrogen (deuterium) distances for the pCA refined by Yamaguchi *et al.*<sup>42</sup> are shown as black (Glu) and green diamonds (Tyr). The correlation plots suggest that short hydrogen bonds are observed between Tyr-42 and the pCA, when the heavy atom distance is < 2.5 Å.

Table 2. Excitation energies ( $\lambda_{\text{max}}$ , in nm) for the four early photocycle intermediates. The table shows the calculated absorption maxima for the protein-embedded and isolated pCA and the respective contributions of strain and electrostatic-induced optical shifts (in eV). The experimental  $\lambda_{\text{max}}$  values were obtained by fitting two Gaussians to the experimental data<sup>9</sup>.

		protein (nm)	Experiment (nm)	isolated (nm)	strain (eV)	electrostatics (eV)
pG	TDDFT	446	452	458	0.12	0.08
	CC2	428		465	0.18	0.13
pR <sub>0</sub>	TDDFT	485	506	512	0.41	0.16
	CC2	483		544	0.61	0.18
pR <sub>1</sub>	TDDFT	472	506	491	0.29	0.12
	CC2	437		501	0.37	0.16
pR <sub>2</sub>	TDDFT	460	464	475	0.19	0.10
	CC2	443		479	0.26	0.06

### Optical Properties of PYP Photocycle Intermediates

Figure 5 shows the dynamically averaged absorption spectra computed at the TDDFT and CC2 levels of theory for the pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> states. The computed absorption maxima are compared to experimental spectra,<sup>9</sup> summarized in Table 2. We find an excellent agreement between the experimental and computational data. The expected absolute error of the employed electronic structure methods is around 0.2 eV (or 40 nm in the 450-500 nm regime) based on benchmarking calculations, which may lead to small discrepancies between the first-principles computed and measured spectra. The quantitative agreement allows us to computationally characterize the likely reason for the observed optical shifts. The dynamically averaged absorption spectra obtained for the pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> states are shown in Figure 6 relative to the absorption properties of the pCA chromophore in gas-phase (dotted line) at the TDDFT/CAM-B3LYP and RVS/CC2 levels of theory. To separate strain and electrostatic polarization effects, we computed the absorption properties of the pCA chromophore in its protein conformation by removing the electrostatically polarizing protein environment around the chromophore (red, isolated), and by computing the absorption properties of the chromophore embedded in its native protein environment (blue, protein).

The TDDFT (CC2) calculations predict that the absorption spectrum of the pG state at 446 nm (430 nm) is redshifted by 0.12 eV (0.18 eV) due to strain effects relative to the gas-phase absorption of the *trans*-pCA at 460 nm.<sup>46</sup> The effect is counterbalanced by a 0.08 eV (0.13 eV) blueshift due to electrostatic stabilization of the surrounding protein environment. Formation of the pR<sub>0</sub> intermediate leads to a large strain-induced redshifted of 0.41 eV (0.61 eV), consistently with the destabilization of the electronic ground state, discussed above. The large strain effect is somewhat counter balanced by the blue-shifting effect of 0.16 eV (0.18 eV) by the protein environment, resulting in absorption peak at around 485 nm (483 nm). Relaxation of this strain energy by 0.1-0.2 eV leads to a gradual blueshift of the predicted spectra in pR<sub>1</sub> and pR<sub>2</sub>.

Interestingly, we find that the partial proton transfer between Tyr-42 and pCA leads to a small population of blueshifted intermediates in the pR<sub>0</sub> and pR<sub>1</sub> states, an effect that becomes more pronounced when the pCA chromophore relaxes. This effect might arise from the increased oscillator strength with decreasing chromophore strain, as suggested by model calculation on the isolated pCA chromophore.

We also computed the absorption spectrum of the 5 ps QM/MM simulation of the I<sub>T</sub> model by Jung *et al.*<sup>16</sup> As expected from the similar conformational dynamics (Figure S4), we also obtain nearly identical spectra for the I<sub>T</sub> and pR<sub>0</sub> trajectories (Figure S5), further supporting that the early 150-ps intermediate leads to large redshift due to strain-induced effects in the pCA.

Our calculations suggest that the molecular strain redshifts all intermediates, while electrostatic polarization has a blueshifting effect on the absorption properties. However, we observe that in contrast to the electrostatic effect, which is 0.08-0.16 eV in all intermediates, the main spectral shift of the pCA chromophore originates from the molecular strain effects, which has a maximum in the recently characterized pR<sub>0</sub> state.

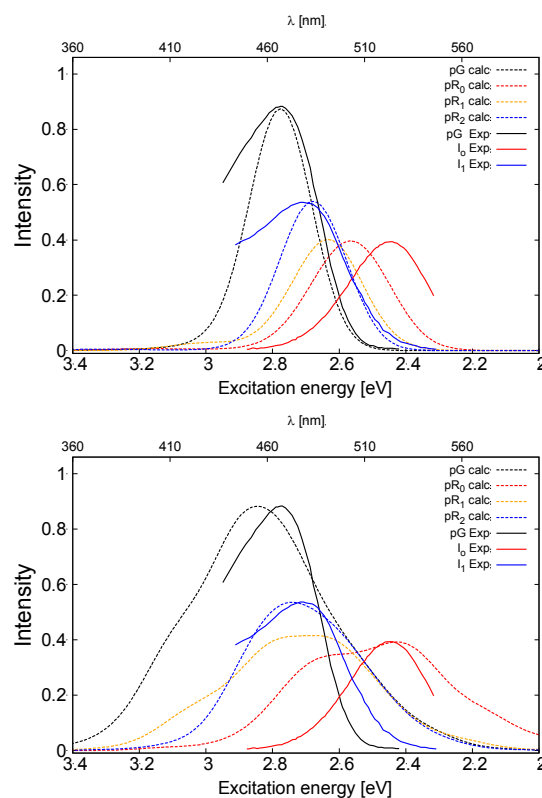


Figure 5. Comparison of the calculated and experimental absorption spectra obtained from Ref.<sup>9</sup> of the PYP photocycle intermediates. TDDFT (CAM-B3LYP/def2-SVP) and CC2 (RVS/CC2/def2-TZVP) values are shown in the upper and lower figures, respectively.

### Conclusions

In conclusion, we have studied here the structure and dynamics of early photocycle intermediates of PYP, and identified how



strain energy, originating from light-absorption is stored in the pCA chromophore using a combination of QM/MM molecular dynamics simulations, and first principles prediction of optical properties. Our data suggest that the recently identified pR<sub>0</sub> state is stable on picosecond timescales, and that the highly strained species is stabilized by hydrogen bonds to the protein environment, which allows the intermediate to store *ca.* 1/3 of the photon energy as molecular strain. Optically, we find that this strained intermediate leads to formation of redshifted species, followed by gradual blueshifting of the spectra with relaxation of the strain. Our simulations also identified short hydrogen bonds and partial proton transfer between Tyr-42 and the pCA chromophore, leading to a small population of blueshifted species in both the computed and experimental optical spectra. While reproducing well the observed optical shifts, our work cannot be used to unambiguously assign the structures responsible for the early steps of the PYP photocycle (refs.<sup>10, 11</sup> and <sup>16, 17</sup>) due to the intrinsic computational errors involved in the employed computational methodology. Our work, nevertheless, clearly shows how the strain effects contribute to the optical tuning in the early PYP photocycle intermediates, providing a powerful methodology to study elementary energy conversion steps in the photocycle of PYP.

Figure 6. The computed ensemble averaged spectra of the four photocycle intermediates pG, pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub> obtained from hybrid QM/MM MD simulations. The absorption maximum of the isolated and protein embedded chromophore are shown in red and blue, respectively, indicating strain (arrow and red values in eV) and electrostatic (arrow and blue values in eV) induced optical shifts. The vertical dotted line shows the averaged absorption maximum for pCA in the *trans* (pG), and *cis* (pR<sub>0</sub>, pR<sub>1</sub>, and pR<sub>2</sub>) conformations. The CAM-B3LYP/def2-SVP spectra are computed as a sum of individual Gaussians, weighted with their respective oscillator strength (solid lines) and as histograms of the VEEs obtained from the MD simulations (vertical bars). The spectra obtained at the RVS/CC2/def2-TZVP level is shown in transparent purple in the background.

## Acknowledgements

We thank Dr. Philip Anfinrud, Dr. Friedrich Schotte and Dr. Hyun-Sun Cho at National Institutes of Health (NIH) for stimulating discussion on the PYP photocycle. The Finnish IT Center for Science and the Leibniz Supercomputing Centre (LRZ). We thank the Jane and Aatos Erkkö foundations for financial support.

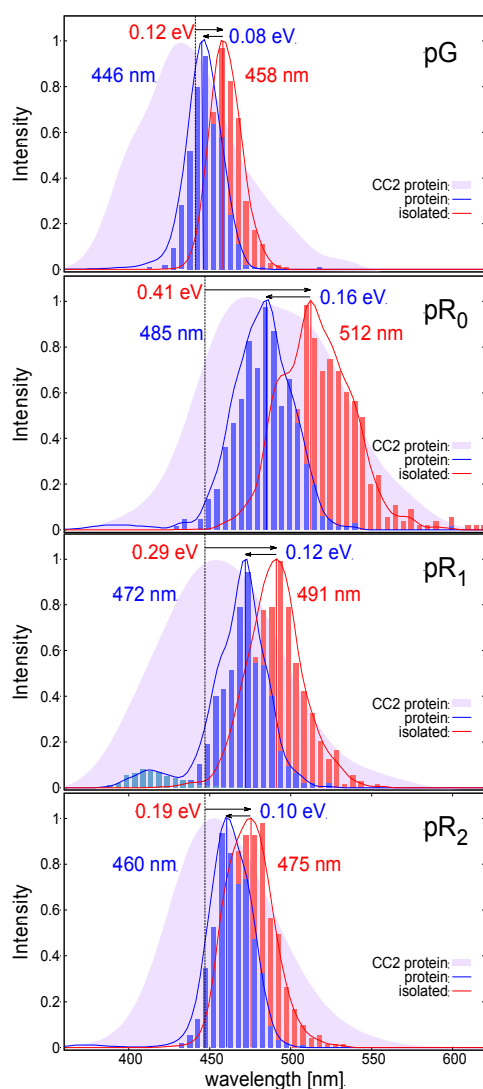
## Abbreviations

pCA • p-coumaric acid, PYP • photoactive yellow protein, CC2 • second-order approximate coupled cluster, RVS • restricted virtual space, QM/MM • Quantum Mechanics/Molecular Mechanics, GS • ground state, ES • excited state, VEE • vertical excitation energy .

## Notes and references

Department Chemie, Technische Universität München (TUM), Lichtenbergstraße 4, D-85747 Garching, Germany. Email: [ana.gamiz@tum.de](mailto:ana.gamiz@tum.de), [vile.kaila@ch.tum.de](mailto:vile.kaila@ch.tum.de)

† Electronic Supplementary Information (ESI) available: Figures S1-S5 and Movies S1 and S2. See DOI: 10.1039/b000000x/



1. T. E. Meyer, *Biochim. Biophys. Acta*, 1985, **806**, 175-183.
2. E. Hustede, M. Liebergesell and H. G. Schlegel, *Photochem. Photobiol.*, 1989, **50**, 809-815.
3. W. W. Sprenger, W. D. Hoff, J. P. Armitage and K. J. Hellingwerf, *J. Bacteriol.*, 1993, **175**, 3096-3104.
4. K. J. Hellingwerf, J. Hendriks and T. Gensch, *J. Phys. Chem. A*, 2003, **107**, 1082-1094.
5. T. E. Meyer, E. Yakali, M. A. Cusanovich and G. Tollin, *Biochemistry*, 1987, **26**, 418-423.
6. T. E. Meyer, G. Tollin, J. H. Hazzard and M. A. Cusanovich, *Biophys. J.*, 1989, **56**, 559-564.
7. T. E. Meyer, G. Tollin, T. P. Causgrove, P. Cheng and R. E. Blankenship, *Biophys. J.*, 1991, **59**, 988-991.
8. W. D. Hoff, I. H. van Stokkum, H. J. van Ramesdonk, M. E. van Brederode, A. M. Brouwer, J. C. Fitch, T. E. Meyer, R. van Grondelle and K. J. Hellingwerf, *Biophys. J.*, 1994, **67**, 1691-1705.
9. Ujj L., Devanathan S, Meyer TE, Cusanovich MA, Tollin G and A. GH., *Biophys J.*, 1998, **75** (1), 406-412.
10. F. Schotte, H. S. Cho, V. R. I. Kaila, H. Kamikubo, N. Dashdorj, E. R. Henry, T. J. Graber, R. Henning, M. Wulff, G. Hummer, M. Kataoka and P. A. Anfinrud, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 19256-19261.
11. V. R. I. Kaila, F. Schotte, H. S. Cho, G. Hummer and P. A. Anfinrud, *Nat. Chem.*, 2014, **6**, 258-259.
12. H. Ihee, S. Rajagopal, V. Šrajer, R. Pahl, S. Anderson, M. Schmidt, F. Schotte, P. A. Anfinrud, M. Wulff and K. Moffat, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 7145-7150.

13. U. K. Genick, G. E. O. Borgstahl, K. Ng, Z. Ren, C. Pradervand, P. M. Burke, V. Šrajer, T.-Y. Teng, W. Schildkamp, D. E. McRee, K. Moffat and E. D. Getzoff, *Science*, 1997, **275**, 1471-1475.
14. B. Perman, V. Šrajer, Z. Ren, T.-y. Teng, C. Pradervand, T. Ursby, D. Bourgeois, F. Schotte, M. Wulff, R. Kort, K. Hellingwerf and K. Moffat, *Science*, 1998, **279**, 1946-1950.
15. S. Anderson, S. Crosson and K. Moffat, *Acta Crystallogr. D*, 2004, **60**, 1008-1016.
16. Y. O. Jung, J. H. Lee, J. Kim, M. Schmidt, K. Moffat, V. Šrajer and H. Ihee, *Nat. Chem.*, 2013, **5**, 212-220.
17. Y. O. Jung, J. H. Lee, J. Kim, M. Schmidt, K. Moffat, V. Šrajer and H. Ihee, *Nat. Chem.*, 2014, **6**, 259-260.
18. G. Wald, *Nature*, 1968, **219**, 800-807.
19. M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim and M. Elstner, *J. Phys. Chem. B*, 2005, **109** (8), 3606-3615.
20. P. Altoe, A. Cembran, M. Olivucci and M. Garavelli, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 20172-20177.
21. V. R. I. Kaila, R. Send and D. Sundholm, *Phys. Chem. Chem. Phys.*, 2013, **15**(13), 4491-4495.
22. X. Zhou, D. Sundholm, T. Wesołowski and V. R. I. Kaila, *J. Am. Chem. Soc.*, 2014, **136**, 2723-2726.
23. A. M. Virshup, B. G. Levine and T. J. Martinez, *Theor. Chem. Acc.*, 2014, **133**, 1506.
24. S. Sekharan and K. Morokuma, *J. Am. Chem. Soc.*, 2011, **133**, 4734-4737.
25. C. M. Isborn, A. W. Götz, M. A. Clark, R. C. Walker and T. J. Martinez, *J. Chem. Theory Comput.*, 2012, **8**, 5092-5106.
26. L. González, D. Escudero and L. Serrano-Andrés, *ChemPhysChem*, 2012, **13**, 28-51.
27. C. Filippi, F. Buda, L. Guidoni and A. Sinicropi, *J. Chem. Theory Comput.*, 2012, **8** (1), 112-124.
28. A. P. Gamiz-Hernandez, A. Neycheva, R. Send, D. Sundholm and V. R. I. Kaila, *Angew. Chem. Int. Edit.*, 2015, **n.a.**, n.a.
29. H. M. Senn and W. Thiel, *Angew. Chem. Int. Edit.*, 2009, **48**, 1198-1229.
30. F. Shirin, G. Gerrit and A. Dreuw, *J. Phys. Chem. B*, 2013, **117** (35), 10071-10079.
31. P. Campomanes, M. Neri, B. A. C. Horta, U. F. Röhrig, V. Stefano, T. Ivano and U. Rothlisberger, *J. Am. Chem. Soc.*, 2014, **136** (10), 3842-3851.
32. T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050-8053.
33. J. Neugebauer, *Chem. Phys. Chem.*, 2009, **10** (18), 3148-3173.
34. V. R. I. Kaila, R. Send and D. Sundholm, *Phys. Chem. Chem. Phys.*, 2013, 4491-4495.
35. R. Send, C.-M. Suomivuori, V. R. I. Kaila and D. Sundholm, *J. Phys. Chem. B*, 2015, **119** (7), 2933-2945.
36. K. Sneskov, T. Schwabe, O. Christiansen and J. Kongsted, *Phys Chem Chem Phys*, 2011, **13**, 18551-18560.
37. B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comp. Chem.*, 2009, **30**, 1545-1614.
38. R. Send, V. R. I. Kaila and D. Sundholm, *J. Chem. Phys.*, 2011, **134**, 214114.
39. H. L. Woodcock, M. Hodošček, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer and B. R. Brooks, *J. Comput. Chem.*, 2007, **28**, 1485-1502.
40. Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. Distasio Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, H. J. M., C. Y. Lin, T. Van Voorhis, S. H. Chien and R. P. S. A. Sodt, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3172.
41. R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165.
42. S. Yamaguchi, H. Kamikubo, K. Kurihara, R. Kuroki, N. Niimura, N. Shimizu, Y. Yamazaki and M. Kataoka, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 440-444.
43. K. Saito and H. Ishikita, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 167-172.
44. K. Saito and H. Ishikita, *Biochim. Biophys. Acta*, 2013, **1827**, 387-394.
45. U. Kosinska Eriksson, G. Fischer, R. Friemann, G. Enkavi, E. Tajkhorshid and R. Neutze, *Science*, 2013, **340**, 1346-1349.
46. I. B. Nielsen, S. Boyé-Péronne, M. O. A. El Ghazaly, M. B. Kristensen, S. Brøndsted Nielsen and L. H. Andersen, *Biophys. J.*, 2005, **89**, 2597-2604.