Supporting Information to Bonetti et al., Hydrogen bond switching among flavin and amino acid side chains in the BLUF photoreceptor observed by ultrafast infrared spectroscopy

Materials and Methods

Sample preparation

*Slr*1694-DNA encoding the full-length SLR-Protein (aa 1 – 150, AccNo: D90913, NP_441709), was amplified by PCR from *Synechocystis* sp.PCC6803 DNA (kindly provided by Dr. I. Maldener, Regensburg) and cloned between EcoRI and HindIII sites of the pET28a(+) vector (Novagen). Protein was expressed in *E.coli* strain BL21(DE3)pLysS over night at 18°C, 0.7 mM IPTG. SLR was purified on Ni-NTA-resin (Qiagen) according to the suppliers instruction. The eluate was dialysed 2x against 200 volumes of 10 mM Tris-HCl, 50 mM NaCl pH 8.0 and concentrated by ultrafiltration (Amicon Ultra-15, M.W.C.O. 10'000, Millipore).

Time-Resolved Spectroscopy

The synchroscan streak camera setup was described previously (1). For time-resolved fluorescence experiments, the sample was diluted to an optical density (OD) at 400 nm of 0.6 per centimeter for a total volume of 8 ml. Both H₂O and D₂O buffers were used of 10 mM TRIS, 50 mM NaCl at a pH/pD of 8.0. To avoid multiple laser excitations on the same volume, the sample was loaded in a flow system and circulated by a peristaltic pump. A 1 mm path length flow cuvette was used. Excitation at 400 nm with a pulse

energy of 10 nJ was provided by frequency doubling the output an amplified Ti:sapphire laser system (Coherent Vitesse-Rega) The pulse duration was 100 fs at a repetition rate of 50 kHz. The fluorescence was collected at 90° to the excitation path using achromatic lenses and detected through a sheet polarizer set at the magic angle (54.7°) with a Hamamatsu C5680 synchroscan streak camera and a Chromex 250IS spectrograph. The streak images were recorded on a Hamamatsu C4880 CCD camera, cooled to -55°C. Streak camera images were obtained for three different time ranges of 200 ps, 500 ps and 2 ns.

Femtosecond time-resolved IR absorption spectroscopy was performed on SIr1694 in H₂O and D₂O buffer (10 mM TRIS, 50 mM NaCl, pH/pD of 8.0). The sample was concentrated to OD_{470} = 125/cm and placed in a sample cell formed by two CaF₂ plates separated by a 20 µm teflon for a total volume of 40 µL (0.25 OD). To avoid multiple laser shots on the same volume during the measurements, the sample was moved by a home-built Lissajous scanner. With a dark relaxation time of the Slr1694 BLUF domain of 20 s in D_2O_1 , 6 s in H_2O and a return time of the Lissajous scanner of > 1 minute, the excited sample volume always remained in the dark state. Changes in the IR absorption were monitored over a spectral window from 1775 cm⁻¹ to 1100 cm⁻¹ in the first nanosecond upon the excitation. The IR absorption of the buffers is characterized by strong and broad bands (1650 cm⁻¹ for H₂O and 1210 cm⁻¹ for D₂O); to avoid distortions and artifacts of the recorded spectra, the sample was probed only in D_2O from 1775 cm⁻¹ to 1430 cm⁻¹, and in H₂O from 1475 cm⁻¹ to 1100 cm⁻¹ referred to as D₂O and H₂O regions, respectively. Three separate measurements are performed for each buffer, leading to six measurements in total. Each measured window had a spectral bandwidth of \sim 180 cm⁻¹, and is partially overlapping with the next one. To check the reproducibility of the data, each dataset consisting of 30 scans taking \sim 30 min in total, was measured several times. The integrity of the sample was checked before, during and after the experiments. No sign of degradation was observed.

The experimental setup for femtosecond IR spectroscopy was described previously(2) and consists of a Ti:sapphire amplified laser (Hurricane, Spectra Physics) providing a 800 nm source, with an output of 0.65 mJ, repetition rate at 1 kHz and duration of 85 fs. A portion of the 800 nm light is use to generate the mid-IR probe pulse. The source pumped an optical parametric generator and amplifier with difference frequency generator (TOPAS, Light Conversion, Vilnius, Lithuania) producing tunable output (2.5–10 μ m) with a bandwidth of ~ 200 cm⁻¹. The probe pulse energy on the sample was attenuated to 1 nJ. A second portion of the 800 nm light was sent into a home-built non-collinear optical parametric amplifier (NOPA) to produce the pump pulse at 470 nm with a pulse energy of 450 nJ, chopped at 500 Hz, and was used to excite the sample (spot diameter ~150 microns). The polarization between pump and probe pulses was set at the magic angle (54.7°) using a Berek rotator (Model 5540; New Focus, San Jose, CA). The probe pulse was collimated into the entrance of a spectrograph and dispersed onto a 32-element mercury-cadmium-tellurium detector array, yielding a spectral resolution of 4 cm⁻¹. The pump-probe spectra were measured in a time windows between -16 ps and 1 ns for a total of 135 delay points, with an instrument response function instrument response function (IRF) of ≈ 160 fs

Data Preprocessing

Before analysis, two preprocessing methods were applied to the raw data. The first corrects for the pre-time-zero signal introduced by thermal lens effects by calculating the average pre-time-zero signal at each wavenumber and subtracting it from the data at that wavenumber. Because no reference probe pulse was used, the noise in the measured spectra consisted mainly of so-called baseline noise (i.e., a flat, structureless offset in the spectra), which is easily recognized from a singular vector decomposition of the residual matrix, the difference between the data and the fit (3). The signal-to-noise ratio of the data was enhanced by subtracting the outer product of the first two singular vector pairs of the residual matrix (being structureless in the time domain and smooth in the wavelength domain) from the data, leading to a factor of 2 reduction in the noise. The experimental error affecting the measured spectra is then ~ 30μ OD.

<u>Data Analysis</u>

The time-resolved data can be described in terms of a parametric model in which some parameters, such as those descriptive of the instrument response function (IRF), are wavenumber-dependent, whereas others, such as the lifetime of a certain spectrallydistinct component, underlay the data at all wavenumbers. The presence of parameters that underlay the data at all wavenumbers allow the application of global analysis techniques(3), which model wavenumber-invariant parameters as a function of all available data. The partitioned variable projection algorithm is well-suited to the optimization of model parameters for global analysis models (4). The algorithm has the further advantage of estimating the standard error of parameters estimates, an advantage that is useful in model selection and validation. A compartmental model was used to describe the evolution of the spectrally distinct components in time. Global analysis was then applied to estimate the lifetime and relative concentration of each component at each wavenumber in the data.

The compartmental model applied to the fluorescence data is reported in figure SI2A. It consists of five components (i=1,2,...,5) decaying in parallel. Table 1 tabulates the estimated decay rates and concentrations of these components. Each component is associated with a decay-associated spectrum (DAS) that describes the concentration of the component in the wavelength dimension.

The femtosecond transient absorption data were first globally analyzed using a kinetic model consisting of sequentially interconverting evolution-associated difference spectra (EADS), i.e. $1\rightarrow 2\rightarrow 3\rightarrow ...$ (figure SI2B) in which the arrows indicate successive mono-exponential decays of increasing time constants, which can be regarded as the lifetime of each EADS. The first EADS corresponds to the time-zero difference spectrum. This procedure enables us to clearly visualize the evolution of the (excited and intermediate) states of the system. In general, the EADS may well reflect mixtures of molecular states and for the specific case of Slr1694 it was demonstrated that such is indeed the case: on the ~10 - 100 ps timescale, a mixture of FAD*, reaction intermediate(s) and long-lived product states make up the EADS (5).

To disentangle the contributions by the various molecular species in the spectral evolution, we performed a target analysis of time-resolved data. Target analysis involves the application of a compartmental model (i.e., a specific kinetic scheme) containing microscopic decay rates expressing intercompartmental transitions, and may be used to test detailed hypotheses regarding the underlying kinetics. The spectrum associated with each component in a target analysis applied to difference absorption data is termed species-associated difference spectra (SADS). In contrast to the EADS, the SADS will represent the spectral signature of the pure molecular species and their kinetics after photon absorption. In this way, the reaction mechanism can be assessed in terms of discrete reaction intermediate states.

The compartmental model from ref. (5)(figure 4A) was applied in a target analysis. The model involves seven compartments, with compartments 1-4 are spectrally identical and represent FAD* (so that compartments 1-4 are associated with a single SADS). Only the first three compartments contribute to population of compartment 5, whereas compartment 4 decays entirely to the ground state. Compartment 5 decays mono-exponentially into compartment 6, and compartment 6 decays mono-exponentially to component 7, which does not itself decay in the time scale of the experiment.

The time constants from the global analysis of the time-resolved fluorescence experiment (Fig. 2, Table 1) are different from those of the sequential analysis of ultrafast IR data (Fig SI3, Supporting information, Table 1). These differences arise because in fluorescence, only the decay of FAD* is observed whereas in the ultrafast IR experiment, also the lifetimes of the FAD radicals contribute to the spectral evolution. Because in the ultrafast IR experiments the number of molecular states increases but the signal-to-noise is not better, we cannot fit the latter data with more time constants. Thus, FAD* and radical dynamics become mixed, resulting in a set of effective time constants in the ultrafast IR sequential analysis that is different from that of the time-resolved fluorescence experiment. In the target analysis of the ultrafast IR data, the time constants from the time-resolved fluorescence experiment are directly plugged in the FAD* compartments (Fig. 4 and Table 1), resulting in a good fit. Thus, time-resolved fluorescence and ultrafast IR data are fitted with a consistent set of rate constants.

The ultrafast IR signals before and around zero time delay were affected by perturbed free induction decay (FID) and by cross-phase modulation artifacts (See Fig. 3 and the collected kinetics in Supporting Information). To take into account the cross-phase modulation artifact, a pulse follower at zero delay was used in the global and target analysis. To avoid inclusion of the perturbed FID signals in the global/target analysis, the signals from -10 ps to 200 fs were given a low weight in the fitting process. Note that the 6 ps component is the fastest time constant in the ultrafast IR experiments that is physically interpreted, much slower than decay of the perturbed FID and cross-phase modulation signals.

Comparison of residuals with one or two intermediates

In Gauden et al. (2006) it has been demonstrated that a kinetic scheme with two (instead of one) intermediates results in an appreciable improvement of the fit of the data in the visible. Here we compare the fit of the data in the infrared using a model with one or two intermediates. Representative traces from the measurements in D₂O are shown in Fig.SI5, and when comparing the fit indicated by the dashed lines between the left and right hand sides an appreciable improvement is obvious. Quantitatively, the root mean square error decreases from 43 to 39 μ OD when introducing the second intermediate. Representative traces from the root mean square error decreases from the measurements in H₂O are shown in Fig.SI6, and again an improvement in the quality of fit is visible. Here the root mean square error decreases from 20 to 19 μ OD when introducing the second intermediates.

proposed by Gauden et al. (2006) provides an adequate description of the infrared measurements as well.

Figures



Fig. SI1: Streak camera images taken of the Slr1694 BLUF domain upon 400nm excitation in H_2O (panel A) and D_2O (panel B) buffers. Each image displays the fluorescence intensity as function of time (vertical axis) and wavelength (horizontal axis).



Fig. SI2: Compartmental models applied to analyze (A) time-resolved fluorescence (parallel scheme); (B) femtosecond time-resolved transient absorption data, global analysis (sequential scheme)



Fig. SI3: Evolution-Associated Difference Spectrum (EADS) with associated lifetime, resulting from global analysis of femtosecond time-resolved visible pump mid-IR probe on *Synechocystis* Slr1694 BLUF domain upon excitation at 475 nm in (A) D₂O and (B and C) H₂O buffer.



Fig. SI4: Concentration profile as function of the time delay for : $FAD*_{1-4}$ (black, Q₁ (red), Q₂ (blue) and Slr_{RED} (green) as follow from target analysis of femtosecond timeresolved IR data on *Synechocystis* Slr1694 BLUF domain in H₂O (panel A) and D₂O (panel B).



Fig. SI5: Comparison of residuals with one or two intermediates. Four representative traces from the measurements in D₂O are shown, the dashed lines indicate the fit. Wavenumber is written along the ordinate. Note that the time axis is linear until 10 ps, and logarithmic thereafter. At the left hand a model with a single intermediate has been used, rms error 43 μ OD. At the right hand a model with two intermediates has been used, rms error 39 μ OD.



Fig. SI6: Comparison of residuals with one or two intermediates. Four representative traces from the measurements in H₂O are shown, the dashed lines indicate the fit. Wavenumber is written along the ordinate. Note that the time axis is linear until 10 ps, and logarithmic thereafter. At the left hand a model with a single intermediate has been used, rms error 20 μ OD. At the right hand a model with two intermediates has been used, rms error 19 μ OD.

Figure SI7 (next page): all kinetic traces of the Slr1694 BLUF domain recorded in D_2O buffer (solid line) with the detection frequency indicated along the ordinate axis. The result of the target analysis is indicated with the dashed line.



Figure SI8 (next page): all kinetic traces of the Slr1694 BLUF domain recorded in H_2O buffer (solid line) with the detection frequency indicated along the ordinate axis. The result of the target analysis is indicated with the dashed line.



References

- 1. Gobets, B., I. H. M. van Stokkum, M. Rögner, J. Kruip, E. Schlodder, N. V. Karapetyan, J. P. Dekker, and R. van Grondelle. 2001. Time-resolved fluorescence emission measurements of photosystem I particles of various cyanobacteria: a unified compartmental model. Biophys J 81:407-424.
- Groot, M. L., L. J. G. W. van Wilderen, D. S. Larsen, M. A. van der Horst, I. H. M. van Stokkum, K. J. Hellingwerf, and R. van Grondelle. 2003. Initial steps of signal generation in photoactive yellow protein revealed with femtosecond midinfrared spectroscopy. Biochemistry 42:10054-10059.
- 3. van Stokkum, I. H. M., D. S. Larsen, and R. van Grondelle. 2004. Global and target analysis of time-resolved spectra. Biochim Biophys Acta 1657:82-104.
- 4. Mullen, K. M. and I. H. M. van Stokkum. 2007. TIMP: An R package for modeling multi-way spectroscopic measurements. Journal of Statistical Software 18.
- Gauden, M., I. H. M. van Stokkum, J. M. Key, D. C. Luhrs, R. Van Grondelle, P. Hegemann, and J. T. M. Kennis. 2006. Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor. Proceedings of the National Academy of Sciences of the United States of America 103:10895-10900.