Supporting Information to

Charge Separation is Virtually Irreversible in Photosystem II Core Complexes with Oxidized Primary Quinone Acceptor

C.D. van der Weij – de Wit, J.P. Dekker, R. van Grondelle, I.H.M. van Stokkum

Department of Physics and Astronomy, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

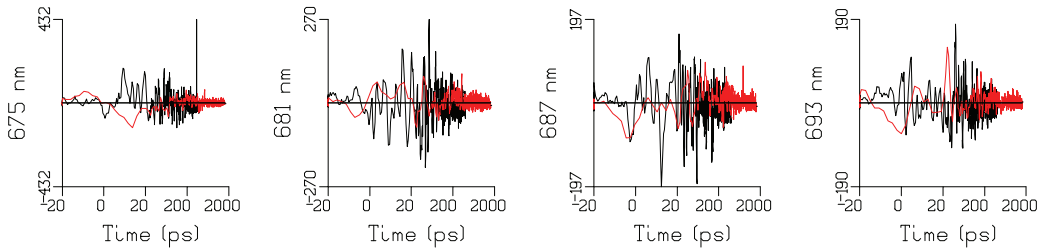

Fig SI.1. Residuals of the fits shown in Figure 2. The traces were measured on a time scale of 0.5 ns (black) and 2 ns (red), with respective instrument responses of 5 and 21 ps FWHM, for wavelengths 675, 681, 687 and 693 nm. Note that the time axis is linear from -20 to 20 ps and logarithmic thereafter.

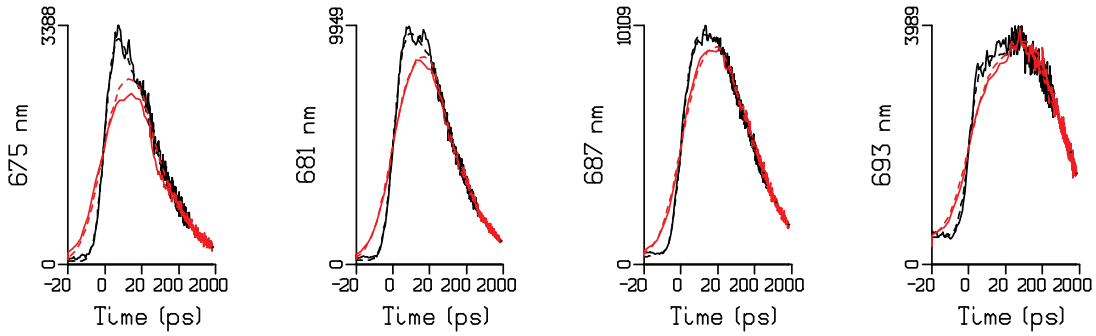

Fig SI.2. Fluorescence time traces (solid) and the global fit (dash) of Figure 8 (77 K) on a time scale of 0.5 ns (black) and 2 ns (red), with respective instrument responses of 10 and 22 ps FWHM, for wavelengths 675, 681, 687 and 693 nm. Note that the time axis is linear from -20 to 20 ps and logarithmic thereafter. Note that "before time zero" long lived red emission (relatively largest at 687 and 693 nm) is present because of the backsweep of the synchroscan streak camera system (see below).

Model based streak data analysis

Because of the limited wavelength resolution of the spectrograph (typically 7 nm FWHM), the averaged streak images can be reduced to a matrix of ≈1000 points in time and 30-60 points in wavelength. The aim of data analysis is to obtain a model-based description of the full data set in terms of a model containing a small number of precisely estimated parameters, of which the rate constants and spectra are the most relevant. With polarized-light experiments also anisotropy parameters come into play. Description of the basic ingredient of kinetic models, the exponential decay, will be given first, followed by a description of how to use these ingredients for global and target analysis[1-3] of the full data. Our main assumption here is that the time and wavelength properties of the system of interest are separable, which means that spectra of species or states are constant. For details on parameter estimation techniques the reader is also referred to[1-4]. Software issues are discussed in[5].

A. Modeling an exponential decay

Here an expression is derived for describing the contribution of an exponentially decaying component to the streak image. The instrument response function (IRF) *i(t)* can usually adequately be modeled with a Gaussian with parameters μ and Δ for, respectively, location and full width at half maximum (FWHM):

$$i(t) = \frac{1}{\tilde{\Delta}\sqrt{2\pi}} \exp(-\log(2)(2(t-\mu)/\Delta)^2)$$

where $\tilde{\Delta} = \Delta/(2\sqrt{2\log(2)})$. The convolution (indicated by an *) of this IRF with an exponential decay (with rate *k*) yields an analytical expression which facilitates the estimation of the IRF parameters μ and Δ:

$$c(t,k,\mu,\Delta) = \exp(-kt) * i(t) = \frac{1}{2}\exp(-kt)\exp(k(\mu + \frac{k\tilde{\Delta}^2}{2}))\{1 + erf(\frac{t-(\mu + k\tilde{\Delta}^2))}{\sqrt{2}\tilde{\Delta}}\}$$

The periodicity of the synchroscan results in detection of the fluorescence that remains after multiples of half the synchroscan period *T* (typically T≈13 ns). Therefore, if lifetimes longer than ~1 ns occur in a sample, the above expression should be extended with a summation over the signal contributions that result from forward and backward

sweeps:

$$c(t,k,T)=\sum_{n=0}^{\infty}e^{-kTn}\{e^{-k(t-\mu+T)}+e^{-k(T/2-t+\mu)}\}=\{e^{-k(t-\mu+T)}+e^{-k(T/2-t+\mu)}\}/(1-e^{-kT})$$

Note that it is assumed here that time zero of the time base corresponds to the zero crossing of the sweep, and that the convolution with the IRF is no longer necessary at times longer than T/2 *(note that the expression in Eq.8 of [3] contains an error).* Adding the previous expressions provides the full model function for an exponential decay recorded with a synchroscan streak camera and will henceforth be denoted by $c^I(k)$:

$$c^I(k)\equiv c(t,k,\mu,\Delta,T)=c(t,k,\mu,\Delta)+c(t,k,T)$$

The simultaneous estimation of up to 5 lifetimes in the range of (sub)ps to ns is common routine.

Because fluorescence samples are relatively dilute, elastic scattering or Raman scattering of the excitation light by water (or of other solvents) can complicate the measurement, if they occur within the analyzed wavelength interval. Such contributions can be modeled with an extra component with a time course identical to the IRF *i(t)*. Usually it is possible to restrict the contribution of scattering to a limited wavelength region.

If the streak image has not been corrected for the instrumental curvature the wavelength dependence of the IRF location μ can be modeled with a polynomial.

$$\mu(\lambda)=\mu_{\lambda_c}+\sum_{j=1}^{j_{\max}}a_j(\lambda-\lambda_c)^j$$

Typically, a parabola is adequate and the order of this polynomial ($j_{\max}$) is two. The reference wavelength λ_c is usually at the center of the streak image.

Sometimes the IRF shape is better described by a superposition of two or even three Gaussians, leading to a superposition description of the exponential decays[4].

B. Global and target analysis

The basis of global analysis is the superposition principle, which states that the measured data $\psi(t,\lambda)$ result from a superposition of the spectral properties $\varepsilon_l(\lambda)$ of the components present in the system of interest weighted by their concentration $c_l(t)$.

$$\psi(t,\lambda) = \sum_{l=1}^{n_{comp}} c_l(t)\varepsilon_l(\lambda)$$

The $c_l(t)$ of all n_{comp} components are described by a compartmental model, that consists of first-order differential equations, with as solution sums of exponential decays. We will consider three types of compartmental models: (1) a model with components decaying monoexponentially in parallel, which yields Decay Associated Spectra (DAS), (2) a sequential model with increasing lifetimes, also called an unbranched unidirectional model[6], giving Evolution Associated Spectra (EAS), and (3) a full compartmental scheme which may include possible branchings and equilibria, yielding Species Associated Spectra (SAS). The latter is most often referred to as target analysis, where the target is the proposed kinetic scheme, including possible spectral assumptions.

(1) with parallelly decaying components the model reads

$$\psi(t,\lambda) = \sum_{l=1}^{n_{comp}} c^I(k_l)DAS_l(\lambda)$$

The DAS thus represent the estimated amplitudes of the above defined exponential decays $c^I(k_l)$. When the system consists of parallelly decaying components the DAS are true species spectra. In all other cases, they are interpreted as a weighted sum (with both positive and negative contributions) of true species spectra.

(2) a sequential model reads

$$\psi(t,\lambda) = \sum_{l=1}^{n_{comp}} c_l^{II} EAS_l(\lambda)$$

where each concentration is a linear combination of the exponential decays,

$c_l^{II} = \sum_{j=1}^{l} b_{jl} c^I(k_l)$, and the amplitudes[6] b_{jl} are given by $b_{11} = 1$ and for $j \le l$:

$$b_{jl} = \prod_{m=1}^{l-1} k_m \Big/ \prod_{n=1,n\neq j}^{l} (k_n - k_j)$$

When the system consists of sequentially decaying components $1 \to 2 \to ... \to n_{comp}$ the EAS are true species spectra. In all other cases, they are interpreted as a weighted sum (with only positive contributions) of true species spectra.

EAS and DAS are directly related[4,7]:

$$DAS_l(\lambda) = \sum_{j=1}^{n_{comp}} b_{lj} EAS_j(\lambda)$$

and inversely:

$$EAS_l(\lambda) = \sum_{j=1}^{n_{comp}} b_{lj}^{-1} DAS_j(\lambda)$$

where the coefficients b_{lj}^{-1} are defined by $b_{1j}^{-1} = 1$ and for $l \le j$:

$$b_{lj}^{-1} = \prod_{m=1}^{l-1} \frac{k_m - k_j}{k_m}$$

In particular, the first EAS is equal to the sum of all DAS. When exciting in the Soret region, the emission at time zero is still zero. Thus the constraint that in the Q_y region the first EAS is equal to zero has been used to estimate the DAS with a subpicosecond lifetime, cf. Figures 1, 5 and 11.

(3) When neither of these two simple models is applicable, a full kinetic scheme may be appropriate. The problem with such a scheme is that while the kinetics are described by microscopic rate constants (their reciprocals are sometimes referred to as *intrinsic* lifetimes), the data only allows for the estimation of decay rates (their reciprocals are referred to as *observed* or *apparent* lifetimes). Thus additional information is required to estimate the microscopic rates, which can be spectral constraints (zero contribution of SAS at certain wavelengths) or spectral relations. This is detailed in[2,4].

Now the model reads

$$\psi(t,\lambda) = \sum_{l=1}^{n_{comp}} c_l^{III} SAS_l(\lambda)$$

where the concentrations c_l^{III} are again linear combinations of the exponential decays, with coefficients that depend upon the microscopic rate constants that describe the transitions between all the compartments. Figures 3, 6 and 9 depict the kinetic schemes applied to our PSII core data. The concentrations of all compartments are collated in a vector $c(t) = \left[c_1(t) \quad c_2(t) \quad \dots \quad c_{n_{comp}}(t) \right]^T$ which obeys the differential equation

$$\frac{d}{dt} c(t) = Kc(t) + j(t)$$

where the transfer matrix K contains off-diagonal elements k_{pq}, representing the microscopic rate constant from compartment p to compartment q. The diagonal elements contain the total decay rates of each compartment. In Figure 3 $c(t) = \left[c_{bA}(t) \quad c_A(t) \quad c_{RC}(t) \quad c_{RP1}(t)\right]^T$ and the input to the compartments is $j(t) = i(t)\left[0.4 \quad 0.4 \quad 0.2 \quad 0\right]^T$. Its K matrix reads:

$$K = \begin{bmatrix} -k_{bA\to A} & k_{A\to bA} & & \\ k_{bA\to A} & -k_{A\to bA} - k_{A\to RC} & k_{RC\to A} & \\ & k_{A\to RC} & -k_{RC\to A} - k_{RC\to RP1} & k_{RP1\to RC} \\ & & k_{RC\to RP1} & -k_{RP1\to RC} - k_{RP1\to final} \end{bmatrix}$$

In Figure 4 (bottom) the c_l^{III} have been drawn, calculated from the estimated parameters, whereas the estimated SAS are shown in Figure 4 (top). Radical Pairs are non-fluorescent and thus their SAS are zero. Note that it has been assumed that the SAS of A and RC are identical, and equal to zero below 668 nm. Because the emission below 668 nm is attributed solely to bA, the multiexponential decay at these wavelengths allows estimation of all forward and backward rate constants. In view of the limited signal to noise ratio, two more spectral constraints[2,4] have been used. The bA SAS is identical to the SAS of A and RC above 710 nm, and between 668 and 710 nm the shapes, but not the amplitudes are assumed to be identical. The thus estimated SAS are considered realistic. These additional spectral constraints enhance the estimability of all forward and backward rate constants. Analogously, in Figure 6 the spectral constraints used are zero SAS of the Radical Pairs and equality of the RC and CP SAS.

Disregarding the energy transfer from the Soret to the Q_y states, and also disregarding the minor 1.4 ns contribution in Figure 1, the global analysis used three lifetimes (12, 49 and 223 ps) and DAS (totalling 3*52=156 spectral parameters at the 52 wavelengths) to describe all data. In the target analysis of Figure 4, only 59 spectral parameters are used instead. The DAS of Figure 5 are fully determined by the SAS of Figure 4A (and thus by these 59 spectral parameters) and the seven microscopic rate constants of Figure 3. This strong reduction from 156 to 59 spectral parameters enables the estimation of more meaningful SAS and kinetic parameters. However, the DAS of Figure 5 are quite different from those in Figure 1.

References

(1) Holzwarth, A. R. Data Analysis of Time-Resolved Measurements. In *Biophysical Techniques in Photosynthesis*; Amesz, J., Hoff, A. J., Eds.; Kluwer: Dordrecht, The Netherlands, 1996; pp 75.

(2) van Stokkum, I. H. M.; Larsen, D. S.; van Grondelle, R. *Biochimica Et Biophysica Acta-Bioenergetics* **2004**, *1657*, 82.

(3) van Stokkum, I. H. M.; van Oort, B.; van Mourik, F.; Gobets, B.; van Amerongen, H. (Sub)-Picosecond Spectral Evolution of Fluorescence Studied with a Synchroscan Streak-Camera System and Target Analysis. In *Biophysical Techniques in Photosynthesis Vol. II*; Aartsma, T. J., Matysik, J., Eds.; Springer: Dordrecht, The Netherlands, 2008; pp 223.

(4) van Stokkum, I. H. M. "Global and Target Analysis of Time-resolved Spectra," Department of Physics and Astronomy, Faculty of Sciences, Vrije Universiteit, Amsterdam, The Netherlands, 2005.

(5) van Stokkum, I. H. M.; Bal, H. E. "A Problem Solving Environment for interactive modelling of multiway data", 2006.

(6) Nagle, J. F.; Parodi, L. A.; Lozier, R. H. *Biophysical Journal* **1982**, *38*, 161.

(7) van Stokkum, I. H. M.; Larsen, D. S.; van Grondelle, R. *Biochimica Et Biophysica Acta* **2004**, *1657*, 82.