ELECTRONIC SUPPORTING INFORMATION

Kinetics and Heterogeneity of Energy Transfer from Light Harvesting Complex II to Photosystem I in the Supercomplex Isolated from Arabidopsis.

Stefano Santabarbara,*[a] Tania Tibiletti,[b] William Remelli[a] and Stefano Caffarri*[b]

[a]Photosynthesis Reserach Unit, Centro di Studio per la Biologia Cellulare e Molecolare delle Piante, Via Celoria 26, 20133 Milan, Italy; [b]Aix Marseille Université, CEA, CNRS, BIAM, Laboratoire de Génétique et Biophysique des Plantes, Marseille 13009, France

APPENDIX S1: *Data analysis using continuous lifetime distributions*

The Exponential Series Method (ESM) is an alternative strategy to analyse excited state kinetics with respect to the most commonly employed linear combination of a few discrete exponential functions. ESM utilises as a model fit-function a parameterised continuous distribution of lifetimes,[1-3] $\rho(\tau)$. The details of the method, both from a theoretical as well as a practical implementation viewpoint, have been previously described.[4, 5] As for the analysis involving a sum of discrete exponentials, it also considers the convolution of the instrument response (IRF) and the decay model functions. The latter, describing the ESM parameterisation of $\rho(\tau)$, has the form: $D_c(t) = \sum_k \alpha_k \cdot \exp(-(t - \delta t)/\tau_k)$, where τ_k are the lifetimes and α_k are the pre-exponential factors, accounting for the area under $\rho(\tau)$ from τ_k to τ_{k+1}. As usual, δt is the time-shift between the decay function and the IRF, which compensates for the excitation pulse jitter. Only α_k are fit parameters whereas the values of τ_k are held fixed. The result of the fit is then an histogram of the amplitudes α_k as a function of τ_k, which represents a discrete approximation of the unknown (*a priori*) distribution $\rho(\tau)$. In the present work, the fit was implemented by initially considering 200 values of τ_k logarithmically spaced in two windows, 1-650 ps and 700-5500 ps, each containing 100 "bins". This allowed to cover the full decay. The maximal number of τ_k "bins" is limited by the number of parameters that the algorithm can handle (i.e. ~250). This might lead to insufficient resolution of closely spaced distributions and/or artificial broadening of the distribution. To obviate to this problem, a refinement step was introduced, by increasing the τ_k bin spacing either in a 1-250 ps or 6-250 ps windows since these represented the most interesting temporal intervals for the samples under investigation, whereas values for τ_k >250 were held fixed from the initial minimisation. The initial guesses were determined from the initial global fit search considering discrete exponentials. These were implemented in the form of normal distributed values centred at the global fit solutions, having widths scaled to the confidence error thereof. A random noise proportional to the residuals obtained from the global analysis was added to the initial guesses to diminish their bias.

Decay traces acquired at different emission wavelengths were fitted independently. The algorithm minimised either χ^2 or the maximum likelihood equivalent χ^2_{ML}.[6] The fit solutions were, within the confidence level, independent from the choice of statistical indicator.

APPENDIX S2: *Spectral-kinetic simulations of excited state relaxation in PSI-LHCI and PSI-LHCI-LHCII*

Excited state relaxation kinetics were simulated using a compartmental model, consisting in a few discrete functional units, whose temporal evolution is described by a system of linear differential equations (ODE) that in matrix notation has the form: $\dot{C}(t) = \mathbf{K} \cdot C(t)$. $C(t)$ and $\dot{C}(t)$ are vectors having as their element the *i*-th compartment population time evolution and their first derivative, respectively. $\mathbf{K}$ is the so-called "rate matrix", *i.e.* a square matrix, containing all the relevant rate constants being either pair-wise singlet energy transfer rates ($k_{i \to j}$ and $k_{j \to i}$) amongst antenna compartments, or the photochemical charge separation rate (k_{tr}) for the reaction centre compartment. Each compartment is characterised by i) the number of chromophores (n_i) and ii) its mean energy ($\bar{E}_i$). The forward and backward energy transfer rates for a pair of compartments are constrained by the Boltzmann distribution (detail balance): $k_{i \to j} / k_{j \to i} = (n_i / n_j) \cdot \exp(-(\bar{E}_i - \bar{E}_j)/k_B T)$, where k_B is the Boltzmann constant and T is the temperature. De-excitation by radiative and non-radiative processes is accounted by an output, described by a constant $k_d = 0.5$ ns^{-1} which is equal for all compartments. The solution of this ODE system is a sum of exponentials of the form: $C(t) = \sum_i V_i \cdot \exp(t\zeta_i)$ where, $\zeta_i = -1/\tau_i$ are the system eigenvalues and V_i are the associated eigenvectors, determined for a given initial condition at t_0. The initial conditions are defined by the excitation of the different compartments by the laser pulse. The values of the eigenvectors define the population (molar fraction) temporal evolution, and do not depend on the detection wavelength. In order to simulate the measured DAS, a spectral function, $G_i(\lambda)$, normalised to unit area, was used for each compartment. $G_i(\lambda)$, which correspond to the Species Associated Spectra (SAS) simulated by a linear combination of Gaussian functions. The first moment of the spectral function is equal to the mean energy of the compartment ($\bar{E}_i$). The spectral-temporal evolution is then given by $C_i(t, \lambda) = C_i(t) \cdot G_i(\lambda)$. The DAS are obtained multiplying the *i*-th element of V_i with $G_i(\lambda)$, which is the equivalent of multiplying a matrix whose rows are the transposed V_i vector, for a column vector composed of the spectral functions $G_i(\lambda)$.

To limit the number of adjustable parameters, the simulations were run considering only the forward energy transfer rates $k_{i \to j}$, the SAS (that define the mean compartment energies ($\bar{E}_i$)) and the macroscopic photochemical trapping rate k_{tr} as variables. The backward energy transfer rates are derived from detailed balance and therefore constrained.

APPENDIX S3: *Representative full time-wavelength excited state dynamics*

Figure S1. The contour plots of representative TCSPC time-wavelength datasets recorded in PSI-LHCI (**A**) and PSI-LHCI-LHCII (**B**) upon unselective excitation at 632 nm are shown.
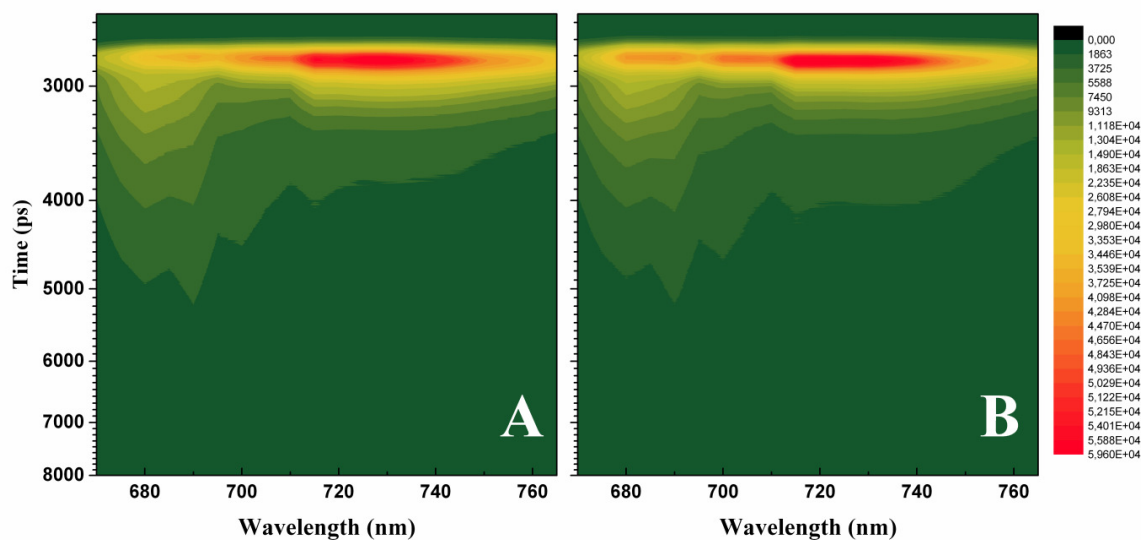
In the presented datasets the emission was monitored between 670 and 760 nm, in 5 nm steps, and each trace was acquired for equal time (900 sec) in order to retain the spectral information. Note that the excitation peak, given by the convolution with the IRF, is at about 2.7 ns. Time traces were corrected for the wavelength-dependent sensitivity of the detector.

Time slices of the contour maps, describing the excited state kinetics are shown for PSI-LHCI in Figure S2 and for PSI-LHCI-LHCII in Figure S3.
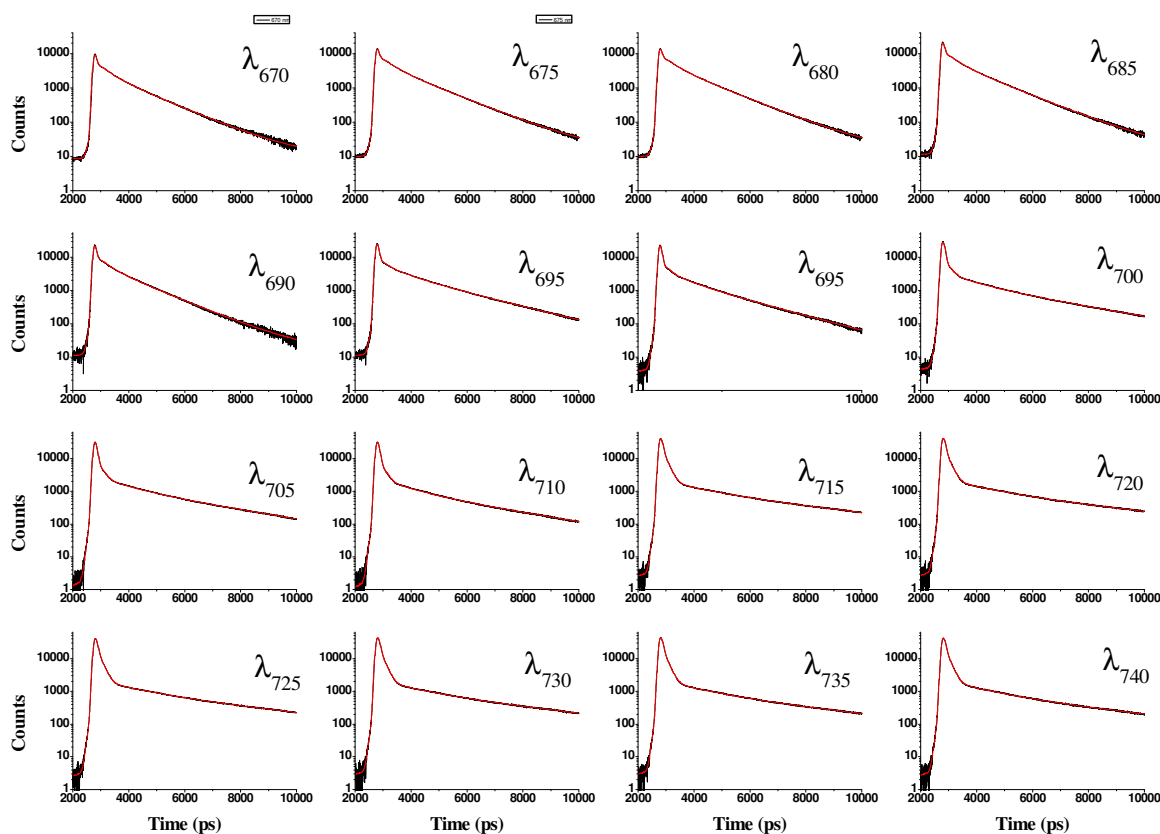
Figure S2. Representative excited state relaxation kinetics of PSI-LHCI supercomplex as a function of the detection wavelength (black lines) together with fits according to the ESM method (red lines).
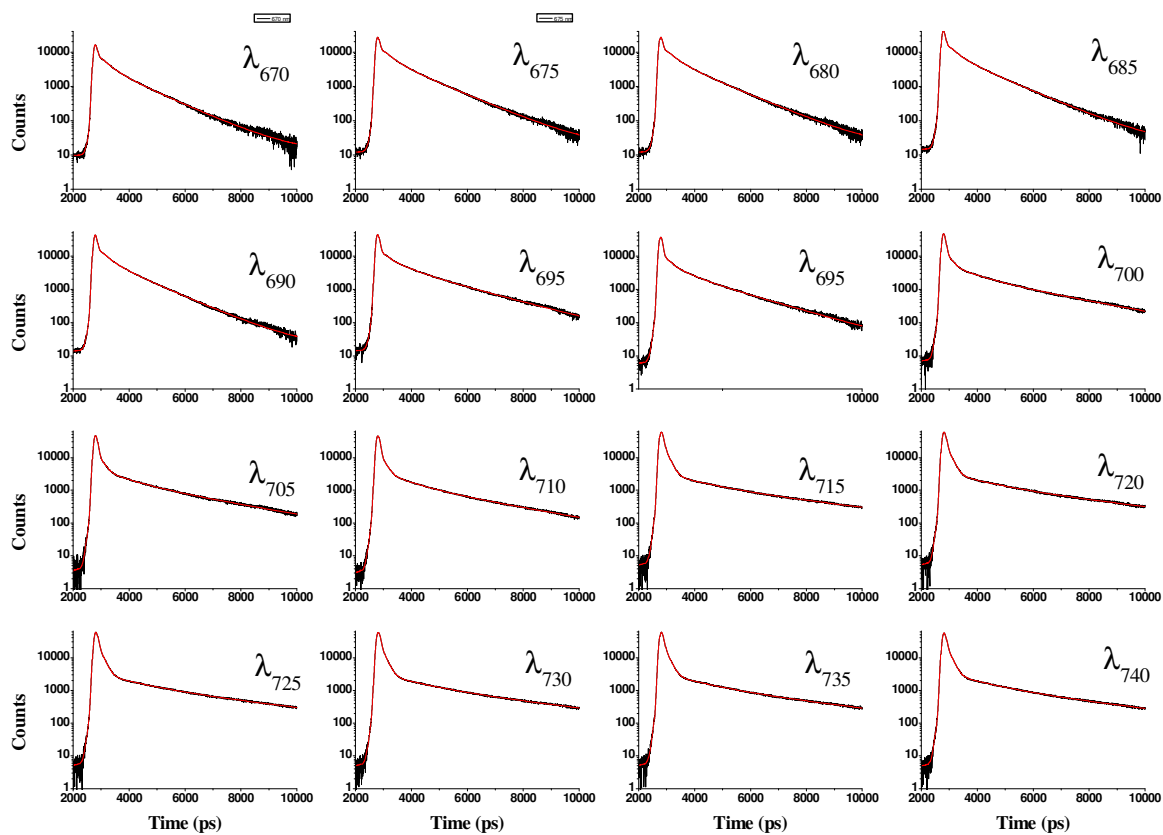

Figure S3. Representative excited state relaxation kinetics of PSI-LHCI-LHCII supercomplex as a function of the detection wavelength (black lines) together with fits according to the ESM method (red lines).

APPENDIX S4: *Global analysis of the excited state decay in PSI-LHCI and PSI-LHCI-LHCII*

Figure S4 shows the results of the global fit analysis of the TCSPC monitored at three selected emission wavelengths (685, 705 and 725 nm) recorded on PSI-LHCI. The results and fits obtained for the PSI-LHCI-LHCII supercomplex are shown in Figure S5. The fits shown in both figures relate to a description in terms of three lifetimes in the sub-nanosecond time window and two having values larger than 1 ns. Considering two long-lived components resulted in an improved description of the time traces tail. However, as it will be discussed below, these components are both of low amplitude and are also attributable to minor purification contaminants. Therefore the attention was focused on the fastest decay lifetimes, which in both supercomplexes were characterised by values of less than 100 ps.

In Figure S4 and Figure S5 are also shown the plots of weighted residuals for the five exponential solutions, from which the fits appear satisfactory. We noticed that the value of the fastest lifetime retrieved from this description displayed some variability, depending on the sample and measurement, when more than 10 independent experimental dataset were compared. Therefore the decays were also investigated by including an additional component (six exponentials fit). This resulted in only a minor improvement in the description when analysing an individual dataset (compare the weighted fit residuals in Figure S4 and Figure S5). However, a general improvement in the stability of the solutions was found amongst different datasets, especially when multiple time-wavelength dataset were fitted simultaneously, and imposing the constraint of equal DAS band-shape, scaled by an amplitude factor amongst datasets. The fastest lifetime, retrieved from the six exponentials fit, fell in the 3-6 ps range, which is rather close to the instrumental resolution limit. The presence of another fast-decaying component of 16-20 ps, which is also close to the instrumental temporal resolution, was always necessary to fit the data. This resulted in relatively large uncertainties on the values of the fastest lifetime and their associated amplitudes in both supercomplexes.

Fitting TCSPC data by iterative convolution methods is a well-known ill-posed mathematical problem.[1, 2, 5, 7, 8] Therefore, different numerical solutions, which diverge only slightly in terms of goodness of fit statistics, can be obtained for each dataset. This is actually what we observed for both descriptions discussed above (five or six exponentials decay functions), albeit with smaller inter-sample variations for the most complex one. Thus in the next paragraphs we present a comparison of the results obtained when considering either three or four lifetimes in the sub-ns range. In both cases the results presented are the average of the solutions found to be common, within the parameters confidence level, for each type of description.

Figure S4 Analysis of the excited state decay of PSI-LHCI detected at selected emission wavelengths (685 nm, 705 nm and 725 nm) by the Global analysis, considering different sets of linear combination of exponential functions. Excitation was at 632 nm (FWHM 3 nm). The central panels show the decay traces of a representative sample (black lines, same as in Figure 2) together with their fits, resulting from a five exponential model function (red lines). The upper panels show the weighted residuals ($w.r_i = (e_i - f_i)/\sqrt{e_i}$; e_i, data point; f_i, fit point), for the five and the six exponentials descriptions. Fit statistics: 685 nm: $\chi^2_{r,5exp}$ = 1.16, $\chi^2_{r,6exp}$ = 1.03; 705 nm: $\chi^2_{r,5exp}$ = 1.05, $\chi^2_{r,6exp}$ = 0.93 and 725 nm: $\chi^2_{r,5exp}$ = 0.95, $\chi^2_{r,6exp}$ = 0.90. Number of counts at the peak channel was: 33299, 18429 and 14178,
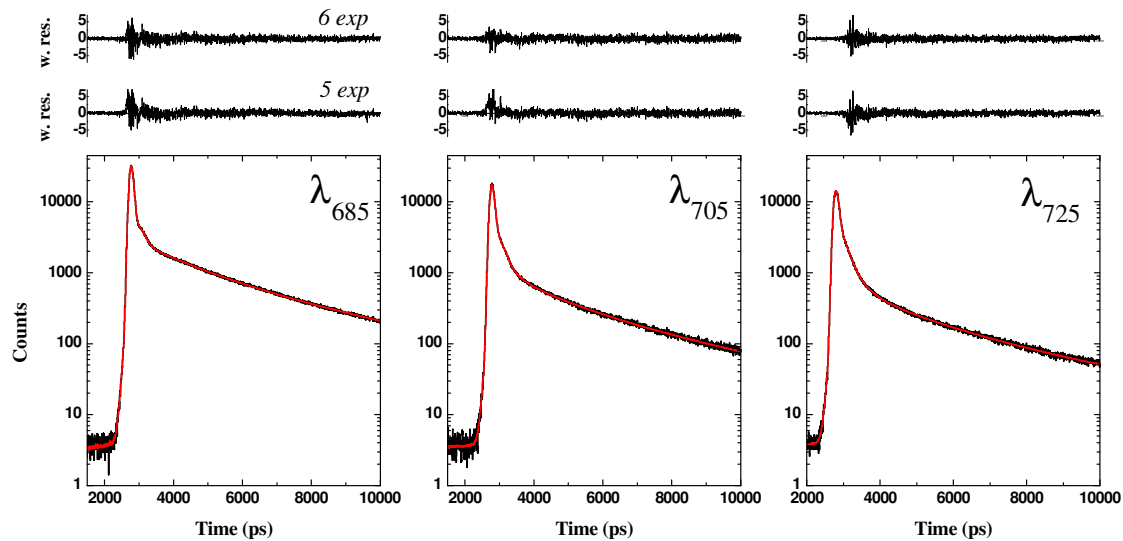
respectively.

Figure S5 Analysis of the excited state decay of PSI-LHCI-LHCII as described for PSI-LHCI in Figure S4. The upper panels show the weighted residuals for the five and the six exponentials descriptions. Fit statistics: 685 nm: $\chi^2_{r,5exp}$ = 1.21, $\chi^2_{r,6exp}$ = 1.11; 705 nm: $\chi^2_{r,5exp}$ = 1.10, $\chi^2_{r,6exp}$ = 1.01 and 725 nm: $\chi^2_{r,5exp}$ = 1.07, $\chi^2_{r,6exp}$ = 0.97. Number of counts at the peak channel was: 46380, 18078 and 14041, respectively.
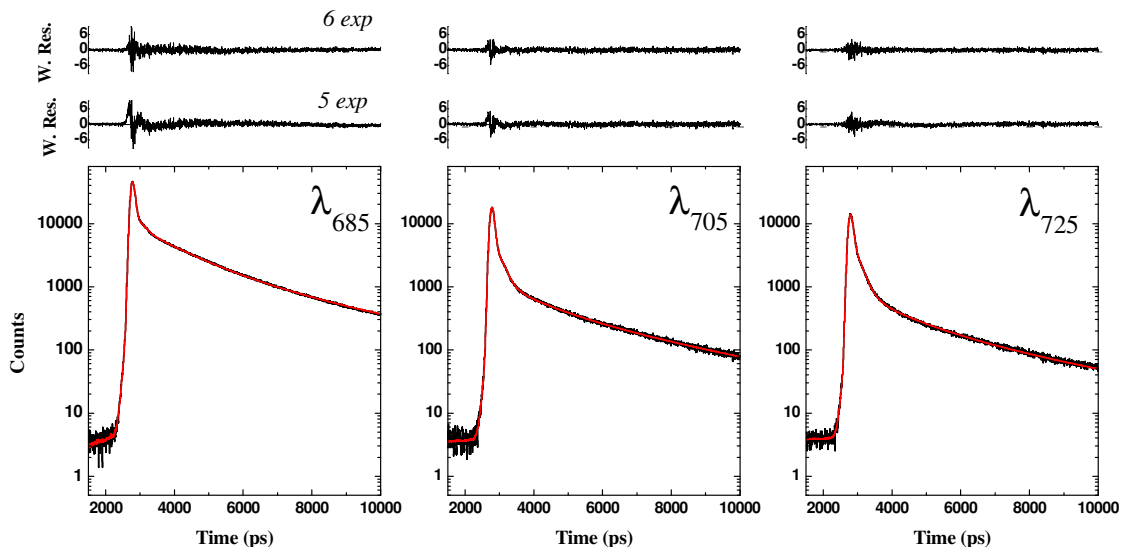
Comparison of the global fit solutions: lifetimes and Decay Associated Spectra (DAS)

Three sub-ns lifetime description. The decay associated spectra (DAS) resulting from the analysis in terms of five exponential functions for PSI-LHCI and PSI-LHCI-LHCII are shown in Figure S6A and Figure S6B respectively. The lifetimes describing the excited state relaxation in PSI-LHCI in the sub-ns window obtained from the global fits are 13±6, 31±7 and 93±8 ps, whereas the two longer lived components are 1.5±0.2 ns and 4.1±0.2 ns. Those describing the decay in PSI-LHCI-LHCII are 16±5, 35±8 and 95±7 ps, whereas the lifetimes of the two slowest decay components are 1.4±0.2 ns and 4.2±0.4 ns. As previously reported,[9, 10] the values of the lifetimes characterising the excited state relaxation in PSI are not significantly affected upon the binding of one LHCII trimer to the supercomplex. Only a moderate increase in the values is observed, which nonetheless remains within the confidence interval of the estimates.

The fastest DAS retrieved from a fit considering three exponential in the sub-ns time window is the most blue-shifted component, having maximal values at 680-685 nm, and showing only weak amplitude above 700 nm, consistently with previous reports on higher plants PSI.[9-12] The most significant difference between the 13±6 DAS of PSI-LHCI and the 16±5 DAS of PSI-LHCI-LHCII, is that the latter shows even lower amplitude in the red emission tail. Both DAS in the 13-16 ps range are dominated by excited state decay and energy trapping, since they have large positive amplitudes at all emission wavelength investigated.

The two remaining lifetimes in the sub-ns time window do not differ significantly in the two supercomplexes. In both cases, the DAS are characterised by rather broad band-shapes and thereby overlap significantly. Yet, even in presence of such broad spectral features, the relative maximum of the ~95 ps DAS appears red-shifted (~735 nm) with respect to the one of the ~35 ps DAS (~725 nm). Similar spectral features were found previously for other plant PSI preparations.[11-14] Although both lifetimes are dominated by long wavelength emission above 700 nm, their band-shapes are asymmetric and extend over the bulk emission in the 670-700 nm spectral window. A specific increase in the intensity of the short-wavelength emission is observed in the 35±8 and the 95±7 ps DAS in PSI-LHCI-LHCII, giving rise to a well-defined secondary maximum centred at 685 nm.
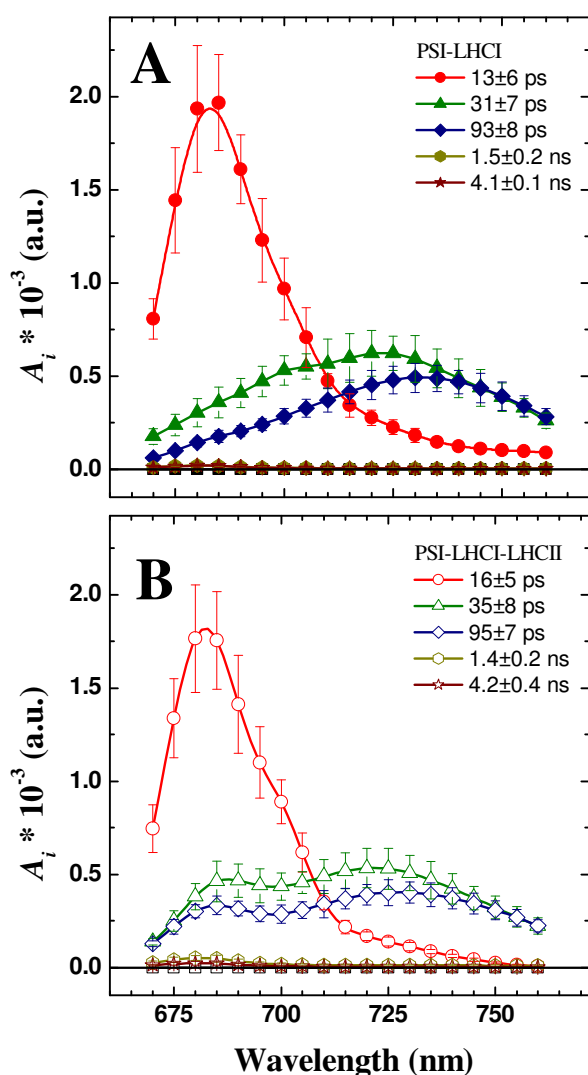

Figure S6 Decay Associated Spectra obtained from the global fits considering five exponential components. **A:** PSI-LHCI. **B:** PSI-LHCI-LHCII. The DAS are normalised to the same total area under the reconstructed steady state emission and are the mean of the results from 10 independent datasets. The confidence levels on amplitudes and lifetimes are from the propagation of uncertainties of the individual fit results.

Four sub-ns lifetime description. The lifetimes describing the excited state relaxation in PSI-LHCI in the sub-ns window are 4.6±1.5, 16±3, 29±8 and 95±9 ps (Figure S7A). Those describing the decay in PSI-LHCI-LHCII are 5.5±1.2, 18.6±3, 32±4 and 97±6 ps (Figure S7B). The two slower decay components are in both samples characterised by lifetimes of 1.1-1.4 ns and 3.9-4.3 ns. These descriptions are, as already discussed above, in general agreement with those previously reported for these preparations[9, 10] except for the fastest lifetime of 3-6 ps, that was not resolved in earlier studies. The values of the two components of ~35 ps and ~95 ps are not significantly altered, within the confidence interval, with respect to the simpler fit description. However, when a faster lifetime, in the order of 3-6 ps, is retrieved from this analysis, the successive lifetime (which was the fastest in the three sub-ns component description) becomes slower in both samples: from 13±6 to 16±3 ps in PSI-LHCI and from 16±5 to 18.6±3 ps in PSI-LHCI-LHCII. This suggests that these two components were observed as a mixed one in the simpler decay description, which is also in line with the observation of rather broad lifetime distributions for $\tau < 20$ ps reported in the main text when the data are analysed by the ESM method.

In PSI-LHCI (Figure S7A) the 4.6±1.5 ps DAS is positive at wavelengths shorter than 705 nm and negative above, indicating that it reflects primarily excited state energy transfer processes. In PSI-LHCI-LHCII (Figure S7B), the mean value of the fastest lifetimes is 5.5±1.2 ps, which, within the parameters confidence limit, is the same as for PSI-LHCI. However, the DAS shows a negative intensity already above 700 nm, leading to a less conservative nature of this component in the larger supercomplex.

The 16±3 ps DAS of PSI-LHCI and the 18.6±3 ps DAS of PSI-LHCI-LHCII, are dominated by a positive feature having maximum at about 685 nm, with an associated shoulder at 700-705 nm and weak intensity above 720 nm. Thus, these components describe excited state decay processes, at all the detected emission wavelengths and in both samples. Similarly to what observed for the simpler description, the intensity in the long wavelength emission region of the 18.6±3 ps DAS of PSI-LHCI-LHCII decreases more steeply than in the 16±3 ps DAS of PSI-LHCI (Figure S7).

The two slower DAS in the sub-ns time window are not modified significantly with respect to those obtained from the solutions considering one less component in this time range, for both PSI-LHCI and PSI-LHCI-LHCII. The band-shape of these DAS also appears to be independent from the exact function used to describe the excited state decay. In fact, these components were relatively well resolved also in the lifetime distributions retrieved by the ESM method (Figures 3-5). It is therefore confirmed that the most obvious spectral difference observed upon the binding of an LHCII trimer to PSI-LHCI is the increase in the relative amplitude of the 32±4 ps and 97±6 ps DAS in 670-690 nm region (Figure S7), which is discernible as distinct secondary peaks, centred at about 685 nm that were not detected in the 29±8 and 95±9 ps DAS of PSI-LHCI.
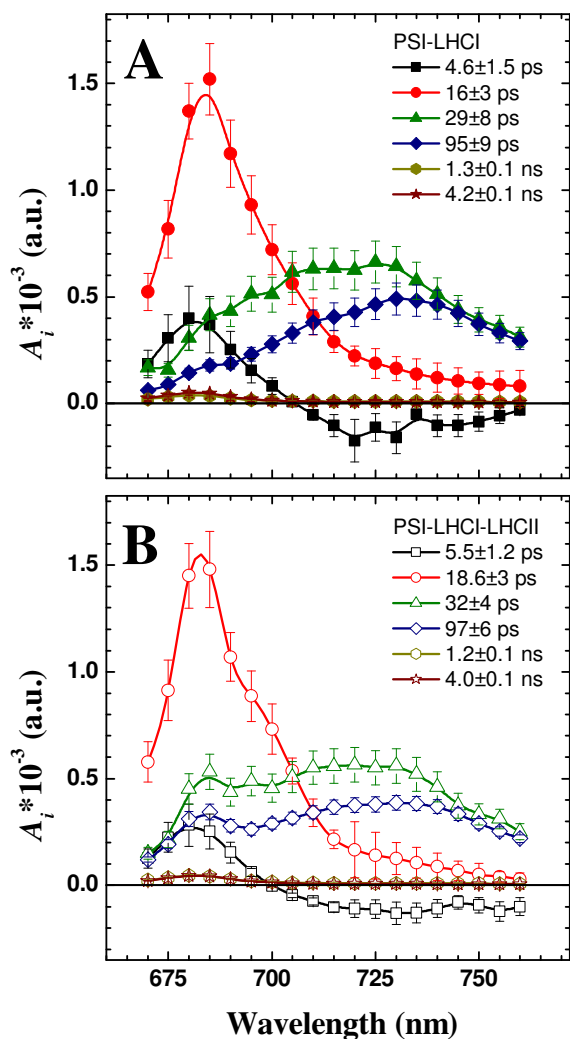
Figure S7 Decay Associated Spectra obtained from the global fits considering six exponential components. **A:** PSI-LHCI. **B:** PSI-LHCI-LHCII. The DAS are normalised to the same total area under the reconstructed steady state emission and are the mean of the results from 10 independent datasets. The confidence levels on amplitudes and amplitude are from the propagation of uncertainties of the singles fit results.

Long-Lived Components. Both long-lived components ($\tau_i > 1$ ns) have weak intensity, irrespectively of the description employed to fit the data (Figures S6 and S7). Their emission is significantly blue-shifted with respect to all the sub-nanosecond lifetimes, having a maximum between 680-685 nm. This becomes evident when calculating their fractional contribution to the steady state emission, which is given by $F_i^{ss}(\lambda) = A_i(\lambda) \cdot \tau_i / \sum_i A_i(\lambda) \cdot \tau_i$. Nevertheless, by virtue of their relatively long lifetimes, they contribute significantly to the steady state emission, irrespectively of their weak fractional amplitudes ($A_i(\lambda) / \sum_i A_i(\lambda)$). This is shown in Figure S8A and S8B, where their reconstructed emission spectra (*i.e.* the sum of their respective fractional contribution) is presented, and compared with that of the short-lived components. In Figure S8A and S8B is also shown the comparison between the steady-state emission recorded under non selective excitation at 435 nm with the reconstruction considering all of the resolved lifetime components, demonstrating a close agreement between time resolved and steady state results. The blue-shifted emission (*e.g.* [15-18]) and the value of their lifetimes, exceeding 1 ns, indicate that the two long-lived and low-amplitude components are due to small fractions of disconnected antenna complexes (*e.g.* [11, 19-21]). Since they likely arise from purification contaminants present in small amounts in both preparations, they can be safely neglected from the analysis of intact supercomplexes.

This is further supported by the inspection of the reconstructed steady-state emission spectra obtained from the sub-ns lifetimes components exclusively (Figure S8C and S8D). The latter resemble closely the spectrum of PSI-LHCI isolated from *Z. mays*,[11-13, 22] which was shown to be free from disconnected antenna contaminants and displayed a maximal emission at 725-730 nm and only a broad shoulder at 685-690 nm. These features were observed in the reconstructed steady state emission of both the PSI-LHCI and PSI-LHCI-LHCII supercomplexes, although the coupling of an LHCII trimer to PSI-LHCI

leads to a relative increase of the intensity at 685 nm with respect to the long wavelength emission tail (Figure S8D). This is due to the larger intensity of the 30-35 ps and 90-95 ps DAS components in PSI-LHCI-LHCII with respect to PSI-LHCI. To a lesser extent, also the moderate increase of the 16 ps lifetime to 18.6 ps (in the six exponentials solution, Figure S7) gives rise to the enhancement of the short-wavelength emission shoulder in the PSI-LHCI-LHCII reconstructed steady-state spectrum as compared with that of PSI-LHCI (Figure S8D).

Figure S8 (C and D) also shows the comparison between the fluorescence emission spectrum reconstructed for $\tau_i < 100$ ps and the steady-state emission recorded upon selective excitation at 715 nm. At this wavelength the absorption cross section of PSI supercomplexes is at least two orders of magnitude larger than that of detached LHCII. The agreement between the steady-state anti-Stokes excitation and the reconstruction of PSI-LHCI(-LHCII) emission from the time-resolved analysis is rather good, particularly above 700 nm where the emission is dominated by the long-wavelength forms [e.g. 12, 22-26], thereby supporting the quality of the kinetics description. It is also interesting to note that the PSI long-wavelength emission band-shape does not appear to be affected by the binding of an LHCII trimer to the supercomplex.
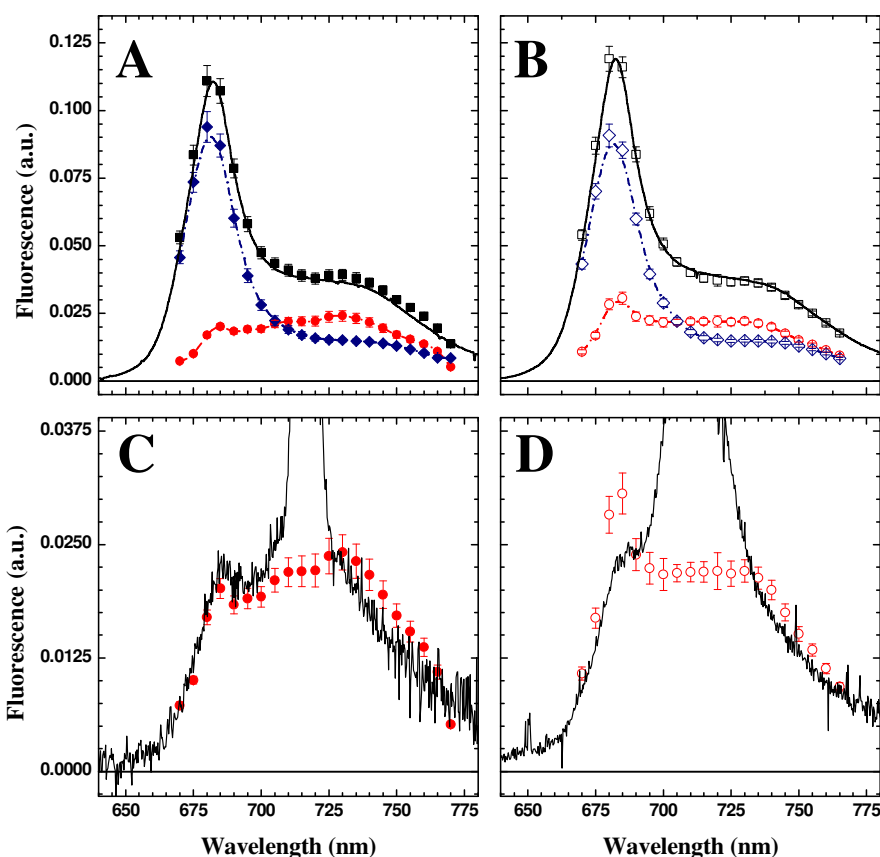

Figure S8 Reconstruction of the steady-state emission spectra from TCSPC data analysis (squares) and comparison with the measured emission spectra upon unselective excitation at 435 nm (solid line) for PSI-LHCI (**A,** solid symbols) and PSI-LHCI-LHCII (**B,** open symbols). The fractional contribution of the fast-decaying ($\tau_i < 100$ ps, red circles) and the long-lived ($\tau_i > 1$ ns, blue diamonds) components for both supercomplexes are also shown. The lower panels show the comparison of the reconstructed steady-state emission from $\tau_i < 100$ ps (red circles), with the steady-state emission detected upon anti-stokes excitation at 715 nm (solid lines) for PSI-LHCI (**C**) and PSI-LHCI-LHCII (**D**).

APPENDIX S5: *Photochemical trapping efficiency.*

The impact of LHCII binding to PSI-LHCI on the photochemical trapping efficiency was estimated by the calculation of the excited state average lifetime which is defined as:

$$\tau_{av}(\lambda) = (\sum_j \tilde{A}_j(\lambda) \cdot \bar{\tau}_j) / \sum_j \tilde{A}_j(\lambda) \qquad\qquad \textit{Equation S1}$$

For the analysis involving the ESM method, when the fitted lifetime distribution $\rho(\tau)$ is deconvolved by a linear combination of asymmetric Gaussian components, the value of the average lifetime (τ_{av}) is evaluated as in Equation 1, but substituting the amplitude A_i, with $\tilde{A}_j$, the area of the j-th asymmetric Gaussian and τ_i with $\bar{\tau}_j$, the mode of the component. This parameter was shown to represent a good approximation for the *effective* trapping time, i.e. $\tau_{av} \sim k_{tr,eff}^{-1}$,[13] where $k_{tr,eff}$ is the macroscopic photochemical rate constant, and it describes the mean time required to form a stable (non-rapidly reversible) radical pair. The spectral dependence of τ_{av} is shown in Figure S9. The parameter was computed either including (Figure S9A) or excluding (Figure S9B) the fastest lifetime density cluster, which, being dominated by energy-transfer character, might lead to bias in the estimation of τ_{av}. In both cases, an increase in the value of τ_{av} is observed for PSI-LHCI-LHCII in the short wavelength emission window, where LHCII emission is expected to contribute the most. On the other hand, in the long wavelength emission window, τ_{av} is little affected by the coupling of LHCII to PSI-LHCI.

The trapping time at the whole photosystem level, $\tilde{\tau}_{av}$, is computed as:

$$\tilde{\tau}_{av} = \sum_u p(\lambda_u) \cdot \tau_{av}(\lambda_u) \qquad\qquad \textit{Equation S2}$$

where $p(\lambda_u)$ is the probability that an excited state resides at a given (u-th) emission wavelength. The values of $p(\lambda_u)$ can be obtained either directly from the steady state emission spectrum normalised per unit area, or, analogously, by computing the steady state emission from time resolved measurements. This is equivalent to the area under the decay traces, which is: $F_{ss}(\lambda) = \sum_i A_i(\lambda) \cdot \tau_i$, or $\sum_j \tilde{A}_j(\lambda) \cdot \bar{\tau}_j$, depending on the analysis approach. The normalisation factor is in all cases: $\sum_u F_{ss}(\lambda_u)$.

The estimated $\tilde{\tau}_{av}$ values were 46±1 ps for PSI-LHCI and 48±1 ps for PSI-LHCI-LHCII, giving photochemical quantum yields of 0.98 for both supercomplexes. Thus, increasing the absorption cross section by more than 20% due to the binding of an LHCII trimer resulted in a slower photochemical trapping observed only at wavelengths shorter than ~700 nm. The effect on the overall quantum yield of charge separation was, however, negligible.
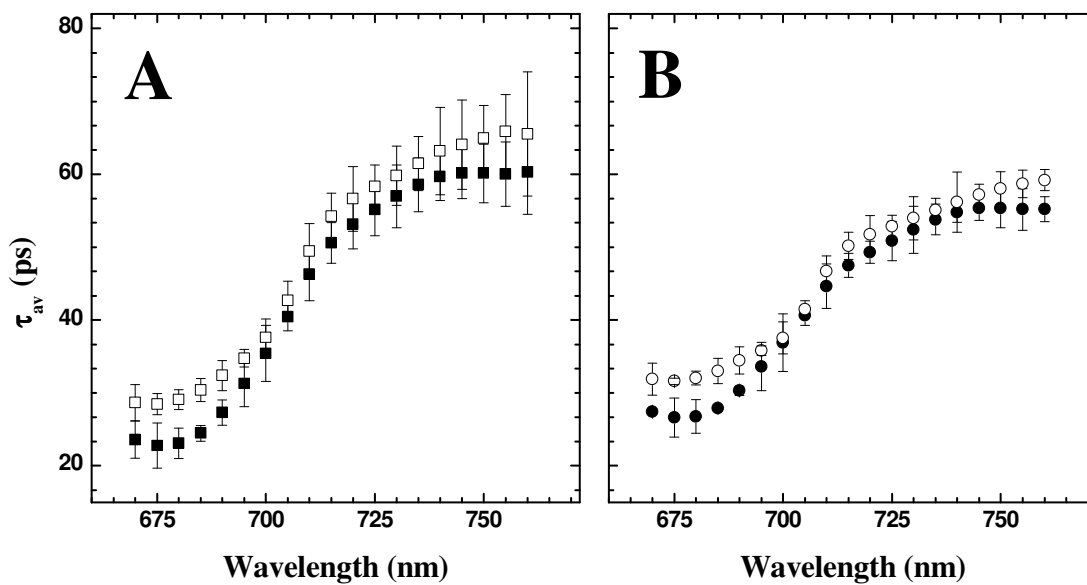
Figure S9. Average excited state lifetimes (τ_{av}) of PSI-LHCI (solid symbols) and PSI-LHCI-LHCII (open symbols), computed from the values retrieved from the ESM-DAS of Figure 6. The values of τ_{av} were computed either considering all $\overline{\tau}_j < 100$ ps (**A,** squares) or excluding the fastest decay component (**B,** circles). Error bars are estimated by the propagation of the uncertainties associated with the values of $\overline{\tau}_j$ and $\tilde{A}_j$.

APPENDIX S6: *Additional information on kinetic modelling*

6.1 "Minimal" kinetic model

   Kinetic model and population evolutions. The *minimal* spectral-temporal model is based on the simplest description of the excited state de-excitation kinetics in terms of three decay components in the sub-ns time window as obtained from the global analysis. Long decay lifetimes are ignored in the model, since they are assigned to uncoupled/detached antenna complexes. The description obtained here for PSI-LHCI (Figure S6) is in excellent agreement with previous analyses of this supercomplex isolated from higher plants.[9-14] Hence the model would apply, qualitatively, also to previous literature data. The minimal model considers then three compartments only. In order to reproduce the experimentally retrieved DAS, two of these compartments need to be assigned to red forms, located in the LHCI complement. The model, illustrated in Figure S10, is already discussed in the main text. Here we present ancillary information.
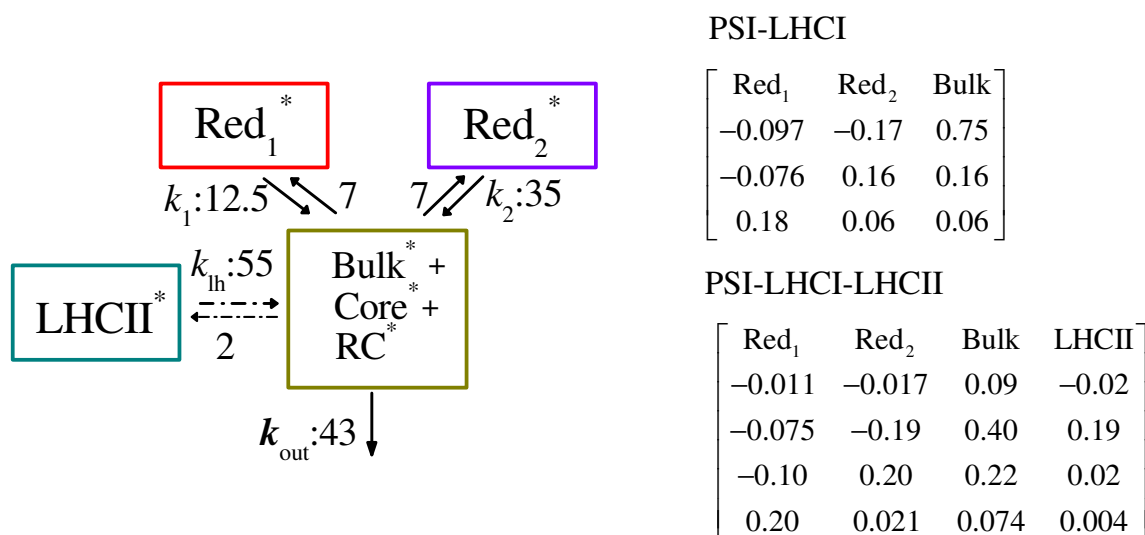

PSI-LHCI

$$\begin{bmatrix} \text{Red}_1 & \text{Red}_2 & \text{Bulk} \\ -0.097 & -0.17 & 0.75 \\ -0.076 & 0.16 & 0.16 \\ 0.18 & 0.06 & 0.06 \end{bmatrix}$$

PSI-LHCI-LHCII

$$\begin{bmatrix} \text{Red}_1 & \text{Red}_2 & \text{Bulk} & \text{LHCII} \\ -0.011 & -0.017 & 0.09 & -0.02 \\ -0.075 & -0.19 & 0.40 & 0.19 \\ -0.10 & 0.20 & 0.22 & 0.02 \\ 0.20 & 0.021 & 0.074 & 0.004 \end{bmatrix}$$

Figure S10 Full set of rate constants (expressed in ns^{-1}) utilised in the minimal model simulations and associated pre-exponential amplitudes (transposed eigenvector matrix) describing each compartment population evolution, considering an unselective excitation in which the initial population is proportional to the number of sites in each compartment. The simulated lifetimes were 14.5, 35, 93 ps for PSI-LHCI and 13.5, 19.5, 36, 95 ps for PSI-LHCI-LHCII.
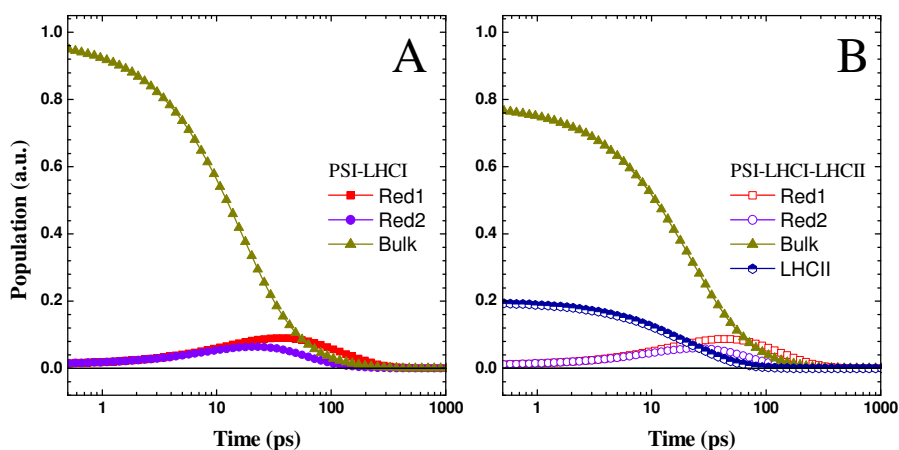

Figure S11 Population evolution of the minimal model compartments for PSI-LHCI (**A**) and PSI-LHCI-LHCII (**B**). Pre-exponential amplitudes are given by the column vectors composing the matrixes presented in Figure S10 whereas the (compartment-independent) lifetimes are listed in the figure legend.

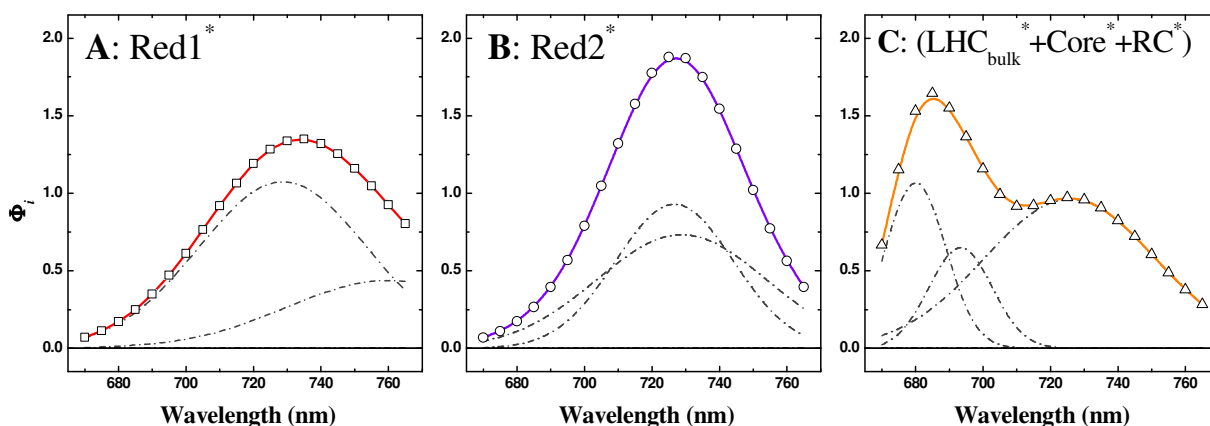*Detailed composition of the SAS spectral functions in minimal model*

Figure S12 Detail of the Gaussian bands utilised to model the Species Associated Spectra (SAS) shown in Figure 8A of the main text. The sub-bands composing each of the simulated SAS are shown as dash-dotted grey lines. Only the composite spectra for the Red_1^* (**A**, red line, open square), Red_2^* (**B** violet line, open circle), LHCI bulk + $Core^*$ + RC^* (**C** orange, open triangle) excited state compartments are shown. $LHCII^*$ was modelled considering a single Gaussian component instead, therefore being identical to the SAS of Figure 8.

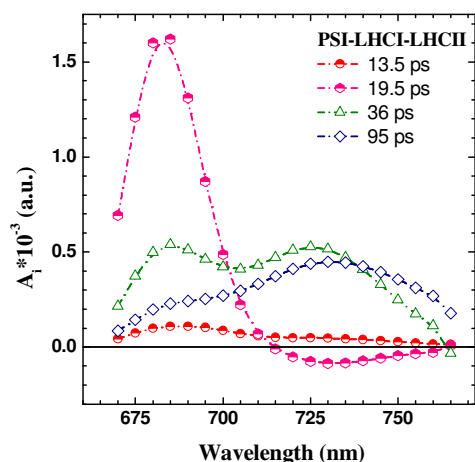*Complete set of simulated Decay Associated Spectra in the minimal model*

Figure S13 Complete sets of the simulated Decay Associated Spectra (DAS) in the minimal kinetic model describing the decay of PSI-LHCI-LHCII. The two fastest DAS (13.5 and 19.5 ps, red and pink) have been pooled for Figure 9B in the main text, in order to match the number of components retrieved from the global analysis (three components in the sub-ns range)..

Simulations considering heterogeneity of energy transfer from LHCII to PSI-LHCI. As discussed in the main text, when considering a homogenous energy transfer rate from LHCII to PSI-LHCI in the order of 35-65 ns^{-1}, the alterations in the experimental DAS observed upon the formation of PSI-LHCI-LHCII can be semi-quantitatively reproduced. However, the relative increase in amplitude of the ~90 ps component in the 675-695 nm emission window is underestimated (whereas that observed of the ~35 nm, is satisfactorily simulated). This led us to consider possible heterogeneity of energy transfer, which is simulated using a linear combination of the model of Figure S10, but using different k_{lh} values. In Figure S14 is shown an example, in which a combination of 0.17:0.83 of slow transferring (k_{lh} =15 ns^{-1}) and fast transferring (k_{lh} =65 ns^{-1}) supercomplex populations results in a satisfactory descriptions of all the principal experimental results, analogously to what presented in Figure 11 of the main

text for the structure-based model. It is worth noticing that considering a slow transferring population only (Figure S14A) does not reproduce the experimental results instead, which is why for an homogenous scenario a faster transfer rate (k_{lh} >30 ns^{-1}) was required.
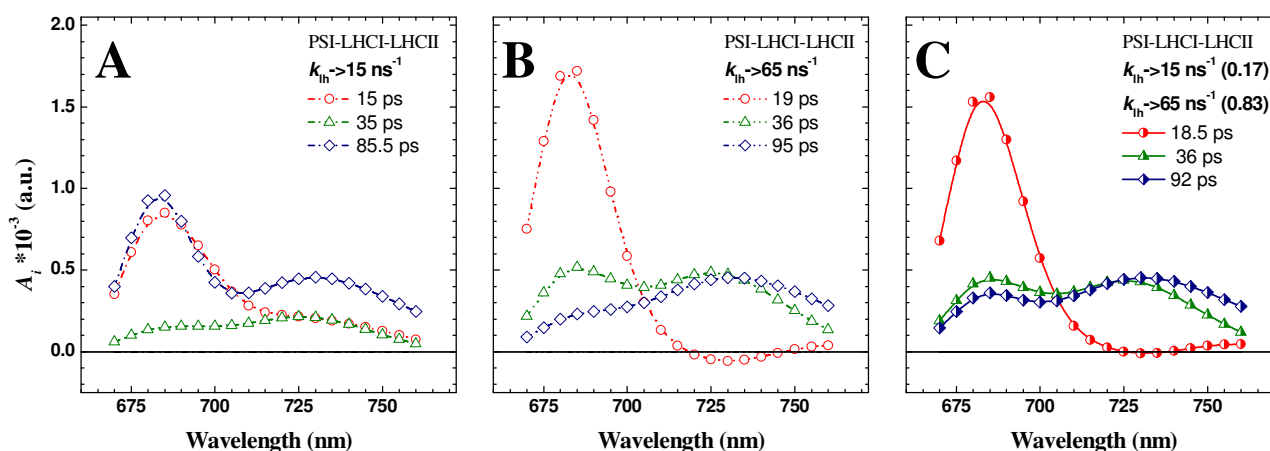

Figure S14 Spectral-kinetic simulations considering heterogeneity of energy transfer from LHCII to the bulk of PSI antenna in the minimal model framework. Simulated Decay Associated Spectra (DAS) of PSI-LHCI-LHCII. **A**: Energy transfer rate k_{lh} = 15 ns^{-1}; DAS: red: 15 ps; green: 35 ps; blue: combined DAS of the simulated 65 ps and 94 ps component (85.5 ps). **B**: Energy transfer rate k_{lh} = 65 ns^{-1}; DAS: red: combined DAS of the simulated 13 ps and 21 ps components (19 ps); green: 36 ps; blue: 95 ps. **C**: DAS resulting from the combination of the two populations shown in panels **A** and **B** in a 0.17 to 0.83 relative ratio on integrated area basis.

6.2 "Intermediate" kinetic model

The number of exponential decays observed in the ESM-DAS (note that this is the same number as in the global analysis for the six exponential fits) defines what we refer to as the "intermediate" model. It therefore considers four compartments for PSI-LHCI (Figure S15). The result of the modelling is given in terms of four DAS and their associated lifetimes, which has to be compared with the deconvolution in Gaussian subcomponents of the ESM retrieved lifetime distributions (ESM-DAS, Figure 6). In order to obtain a semi-quantitative agreement between the simulations and the experimental results, the kinetic/spectral model requires considering two compartments associated with the long-wavelength emitting Chl forms of the LHCI antenna (referred to as the Red$_1$* and Red$_2$* states), a compartment associated with the bulk antenna (Bulk*), comprising both the chromophores coordinated by the core as well as the LHCI pigments other than those described by the Red$_1$ and Red$_2$ states. Moreover, a compartment describing the excited state of the reaction centre (RC*) is to be taken into account. The RC* and the antenna pigments other than red states Red$_1$ and Red$_2$ were considered as a single compartment in the "minimal" model instead. This is the sole difference in between the two model descriptions. On the other hand, the "intermediate" considers one less compartment with respect to the extended structure-based one where, as discussed in the main body, the core and the LHCI bulk antenna are considered as independent compartments.
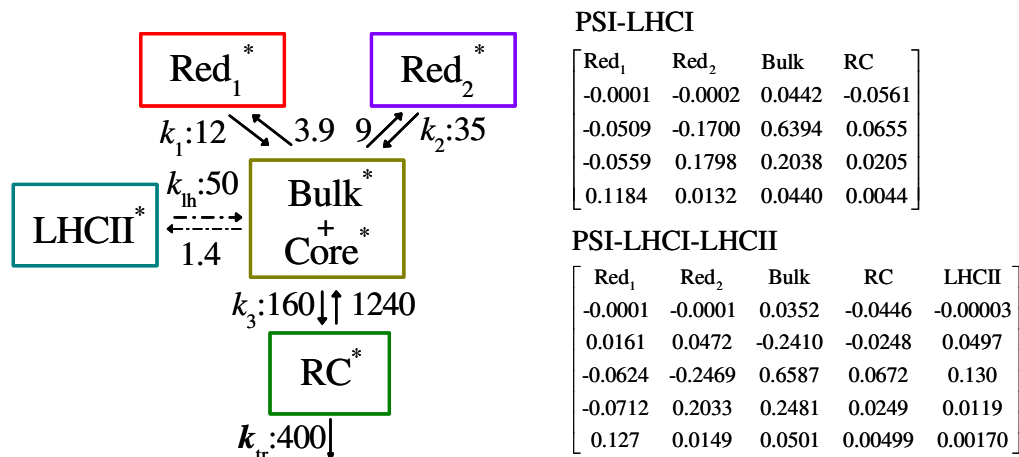
Red$_1^*$ Red$_2^*$

k_1:12   3.9   9 k_2:35

k_{lh}:50   Bulk* + Core*

LHCII* 1.4

k_3:160 ↓↑ 1240

RC*

k_{tr}:400

PSI-LHCI

Red$_1$	Red$_2$	Bulk	RC
-0.0001	-0.0002	0.0442	-0.0561
-0.0509	-0.1700	0.6394	0.0655
-0.0559	0.1798	0.2038	0.0205
0.1184	0.0132	0.0440	0.0044

PSI-LHCI-LHCII

Red$_1$	Red$_2$	Bulk	RC	LHCII
-0.0001	-0.0001	0.0352	-0.0446	-0.00003
0.0161	0.0472	-0.2410	-0.0248	0.0497
-0.0624	-0.2469	0.6587	0.0672	0.130
-0.0712	0.2033	0.2481	0.0249	0.0119
0.127	0.0149	0.0501	0.00499	0.00170

Figure S15 Full set of rate constants (expressed in ns^{-1}) utilised in the intermediate model simulations, yielding the presented pre-exponential amplitudes (transposed eigenvector matrix) describing each compartment population evolution. Since in the experiments the sample is excited at 632 nm, which is largely unselective, the initial condition (excitation vector) was determined assuming fractional scaling according to the number of sites in each compartment. The simulated lifetimes were 0.5, 16, 39, 88 ps for PSI-LHCI and 0.5, 15, 21, 40, 89 ps for PSI-LHCI-LHCII.

The two compartments describing the red-forms have mean energies of $\overline{E}_{red1}$ = 1.69 eV and $\overline{E}_{red2}$ = 1.71 eV, corresponding to photon energies of 734 nm and 726 nm, respectively. The slight difference in energy between the Red$_1$ and Red$_2$ states is determined by the larger bandwidth of the latter, resulting from the superimposition of spectral forms having maximal emission at 735 nm and 715 nm, as shown in the Species Associated Spectra (SAS) of Figure S16A. Different values for the energy transfer rate between the red states and the bulk of the antenna had to be considered. This was also the case in previous modelling of higher plants PSI, as reported by Engelmann *et al.*[11] for *Z. mays* PSI-LHCI and by Croce and coworkers[27] for the *A. thaliana* supercomplexes. The bulk antenna compartment has the broader SAS, since this considers both the emission from core antenna and the LHCI chromophores non-explicitly described by the Red$_1$ and Red$_2$ states. Its mean energy is 1.76 eV, whereas the mean energy of the RC* compartment is set at 1.73 eV, the same value that is discussed for the structure-oriented model. Photochemical charge separation is described in the intermediate model by a macroscopic constant (k_{tr}), which is independent on the detailed mechanisms of charge-separation and charge stabilisation. Considering a value of k_{tr} = 400 ns^{-1} (i.e. $1/k_{tr}$ ~ 2.5 ps), the model yields lifetimes of 0.5, 16, 39 and 88 ps, which are in substantial agreement with the experimental results, but for the value of the fastest lifetime that underestimates the one obtained experimentally. Moreover, the amplitude of the DAS associated with the fastest modelled component is underestimated and the band-shape is somewhat distorted with respect to experimental values (Figure S16B). However, this is the component that is associated with the largest experimental uncertainties and that displays the widest kinetic dispersion. Therefore, we consider that the good agreement between the modelled DAS and lifetimes for the remaining components τ >10 ps indicates that the model is substantially reliable.
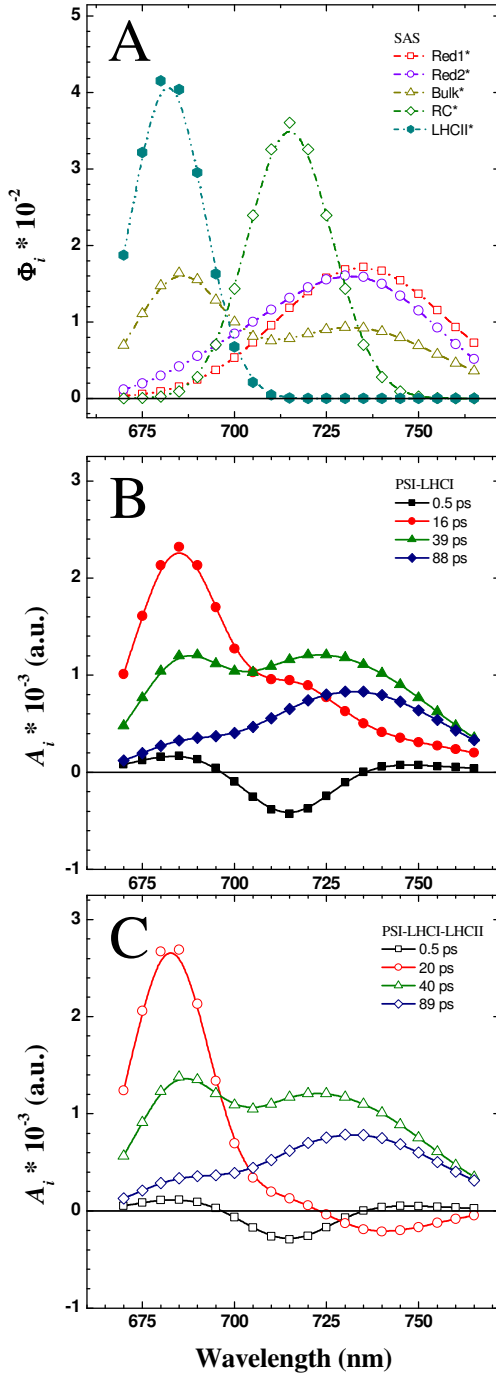
Figure S16 Intermediate kinetic model: Species Associated Spectra (SAS) and simulated Decay Associated Spectra (DAS) describing the excited state relaxation in PSI-LHCI and PSI-LHCI-LHCII. (**A**): SAS. Dark Cyan, solid symbols: LHCII ($\overline{E}$=1.82 eV); dark yellow, open triangles: PSI bulk (Core +LHCI bulk) antenna ($\overline{E}$=1.76 eV); violet, open circles: Red emitting state 2 ($\overline{E}$=1.71 eV); red, open square: Red emitting state 1 ($\overline{E}$=1.69 eV); green, open diamonds: RC* ($\overline{E}$=1.73 eV). Each SAS is normalised to the unit area. (**B**): DAS of PSI-LHCI. Black line: 0.5 ps; red line: 16 ps; green line: 39 ps; blue line: 88 ps. (**C**): DAS of PSI-LHCI-LHCII. Black line: 0.5 ps; red line: combined DAS of the simulated 15 and 21 ps components (20 ps); green line: 40 ps; blue line: 89 ps. The DAS are normalised to yield the same integrated steady state emission.

The formation of the PSI-LHCI-LHCII supercomplex is, in this model, described by the coupling of a compartment having mean energy $\overline{E}_{LHCII}$ = 1.82 eV ($\overline{\lambda}$ =682 nm), with the bulk of the PSI antenna. The energetics of the compartment describing LHCII were derived from previous measurements on the isolated trimer.[15,16,18] The presence of an LHCII compartment in the model gives rise to an additional decay lifetime, that, being unresolved in the experiment, is expected to fall in the τ < 25 ps interval. Moreover, a specific DAS associated to it is not discerned experimentally. In order to obtain a straightforward comparison with the experimental data, the number of the modelled and the detected DAS needs to be the same. This is achieved in the simulations by merging the DAS associated to closely laying lifetimes. This is implemented by the summation of their respective fractional contribution to the steady-state population, weighted by the mean of their lifetimes: $u\sum_{n=1}^{u} A_n(\lambda)\tau_n / \sum_{n=1}^{u} \tau_n$, were u is the number of simulated DAS that are pooled together. The simulations of the individual (un-pooled) DAS are presented in Figure S17. In order to reproduce the changes in the DAS bandwidth observed in the experiments upon the association of LHCII with PSI, it is necessary to consider an energy transfer rate from LHCII to the bulk of PSI antenna in the order of 50 ns^{-1}, from which lifetimes of 0.5, 15, 21, 40 and 89 ps are obtained for PSI-LHCI-LHCII. The compartments population evolutions, for both supercomplexes, are shown in Figure S18. The principal differences in the DAS resulting from the coupling of LHCII are (Figure S16B and C):

i) the change in the band-shape of the $\overline{\tau}$ =20 ps DAS (pooled 15 and 21 ps lifetimes), which has a larger relative intensity in the short-wavelength emission region and decreases more sharply at longer emission

wavelengths, eventually having negative amplitude above 720 nm. The decrease in intensity on the emission tail observed in the experiment is therefore reproduced, but for the simulated DAS displaying negative amplitude in the red tail. Some discrepancies are to be nevertheless expected because of the simplified description of LHCII emission by a single Gaussian band centred at 682 nm, which neglects the vibronic emission tail of the complex that would in part compensate for the negative (energy transfer) amplitude in the simulations.

ii) the broadening of the DAS associated with the two longest components, due to the appearance of a secondary maximum at about 685 nm, which is the other characteristic alteration observed in the experimental EMS-DAS. Although this is qualitatively reproduced in the simulations, it appears more obvious for the 40 ps rather than the 89 ps DAS. This discrepancy remained even when considering the more extended structure based model.

We notice that the macroscopic rate constant of energy transfer between LHCII and the bulk antenna retrieved here (50 ns^{-1}) is comparable to the value obtained either from the structure-based (45 ns^{-1}) or the minimal (60 ns^{-1}) model. In all cases, values in the 35-65 ns^{-1} range yielded results acceptable on semi-quantitative basis. Thus, since we are not seeking for an exact value but for a reasonable estimate, both models converge to rather similar outcomes. This can be taken as an indication of the robustness of the energy transfer estimates, when also their associated confidence intervals are considered.

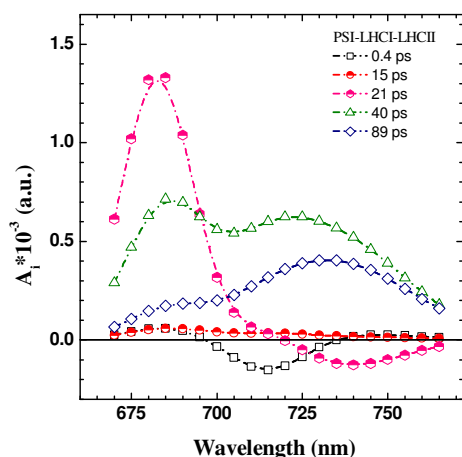*Complete set of simulated Decay Associated Spectra for PSI-LHCI-LHCII*

Figure S17 Complete sets of DAS for PSI-LHCI-LHCII obtained from the intermediate kinetic model. The DAS of the 15 ps, (red circles) and 21 ps (magenta hexagons) that were pooled in order to match the number of experimentally retrieved spectra, are indicated by half-filled symbols. All other symbols as in the legend of Figure S16C.

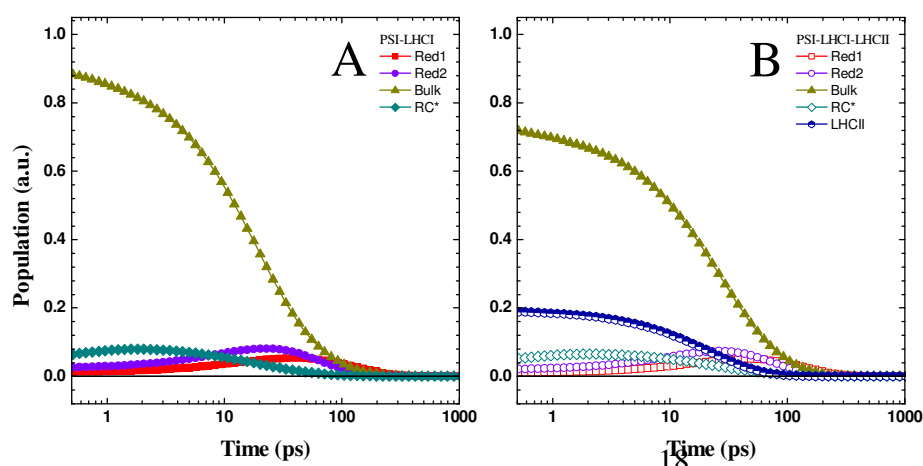*Simulated compartment population evolutions*

Figure S18 Population evolution of the compartments considered in the intermediate model of PSI-LHCI (**A**) and PSI-LHCI-LHCII (**B**). Lifetimes for PSI-LHCI: 0.5 ps, 16 ps, 39 ps, 88 ps; and for PSI-LHCI-LHCII: 0.5 ps, 15 ps, 21 ps, 40 ps, 89 ps. Pre-exponential amplitudes and lifetimes are shown in Figure S15 and its legend, respectively.

6.3 "Structure-based" kinetic model
Kinetic model and population evolutions

PSI-LHCI

Red$_1$	Red$_2$	Bulk	Core	RC
0.00004	0.00005	-0.00611	0.07156	-0.08915
0.00408	0.00586	-0.10856	0.11653	0.03466
-0.06193	-0.15963	0.32083	0.25926	0.06418
-0.08644	0.15778	0.12979	0.08126	0.01955
0.15588	0.01919	0.04196	0.02371	0.00563

PSI-LHCI-LHCII

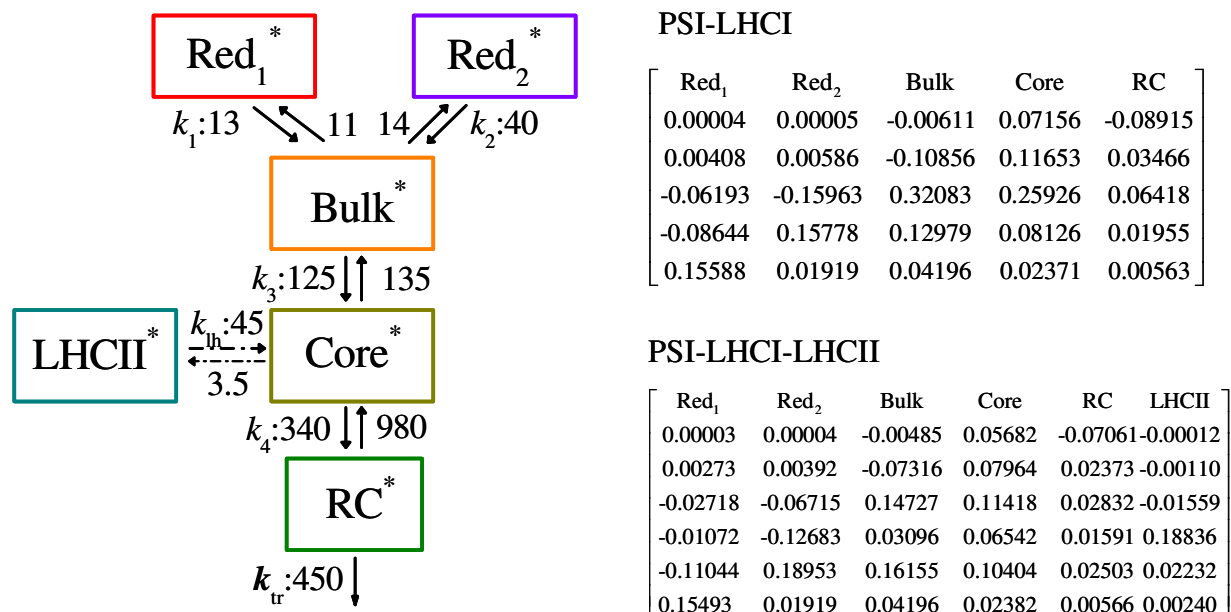| Red$_1$	Red$_2$	Bulk	Core	RC	LHCII
0.00003	0.00004	-0.00485	0.05682	-0.07061	-0.00012
0.00273	0.00392	-0.07316	0.07964	0.02373	-0.00110
-0.02718	-0.06715	0.14727	0.11418	0.02832	-0.01559
-0.01072	-0.12683	0.03096	0.06542	0.01591	0.18836
-0.11044	0.18953	0.16155	0.10404	0.02503	0.02232
0.15493	0.01919	0.04196	0.02382	0.00566	0.00240

Figure S19 Full set of rate constants (expressed in ns^{-1}) utilised in the structure-based model simulations, yielding the following pre-exponential amplitudes (transposed eigenvector matrix) describing each compartment population evolution, considering an unselective excitation in which the initial population is proportional to the number of sites in each compartment. The simulated lifetimes were 0.6, 3.4, 14.5, 34, 93 ps for PSI-LHCI and 0.6, 3.4, 14, 22, 34, 93 ps for PSI-LHCI-LHCII.
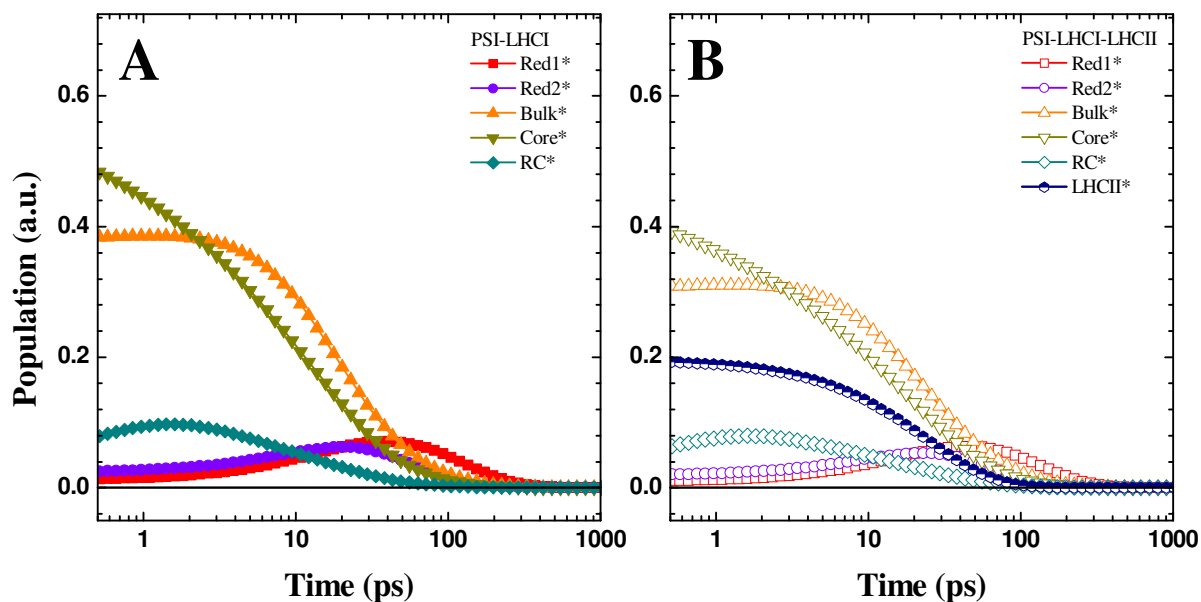

Figure S20 Population evolution of the compartments considered in the structure-based model of PSI-LHCI (**A**) and PSI-LHCI-LHCII (**B**). Pre-exponential amplitudes are given by the column vectors composing the matrixes presented in Figure S19 whereas the (compartment-independent) lifetimes are listed in the figure legend.

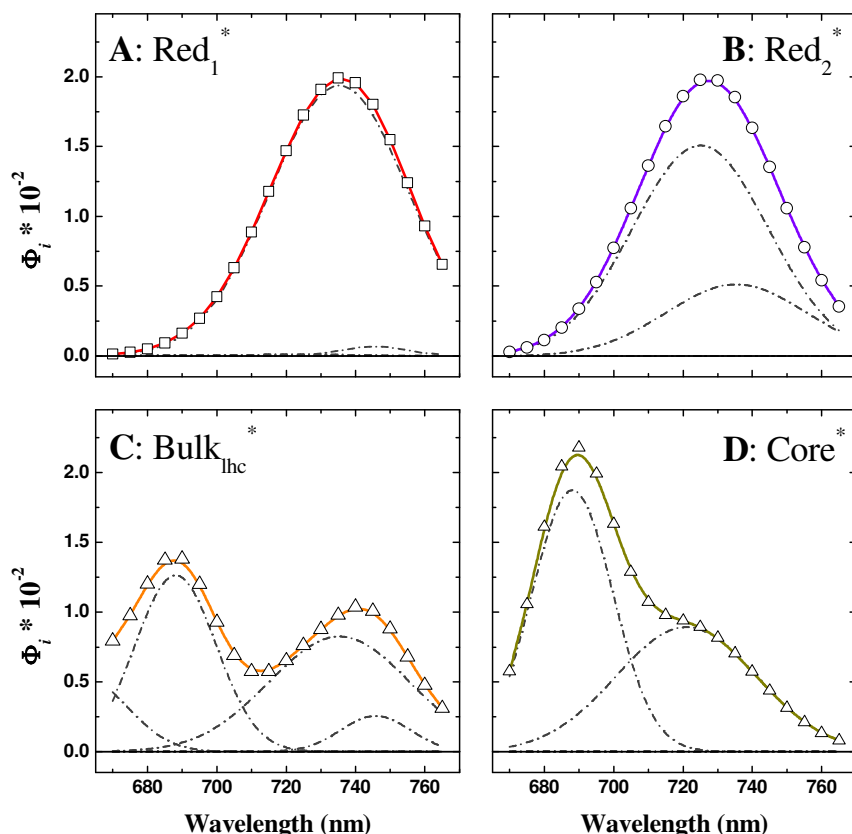### Detailed composition of the SAS spectral functions

Figure S21 Detail of the Gaussian bands utilised to model the Species Associated Spectra (SAS) shown in Figure 8B of the main text. The sub-bands composing each of the simulated SAS are shown as dash-dotted grey lines. Only the composite spectra for the Red_1^* (**A**, red line, open square), Red_2^* (**B** violet line, open circle), LHCI $bulk^*$ (**C** orange, open triangle) and the $Core^*$ (**D** dark yellow, open triangle) excited state compartments are shown. RC^* and $LHCII^*$ were modelled considering a single Gaussian component, therefore being identical to the SAS of Figure 8B.

Complete set of simulated Decay Associated Spectra in the structure-based model
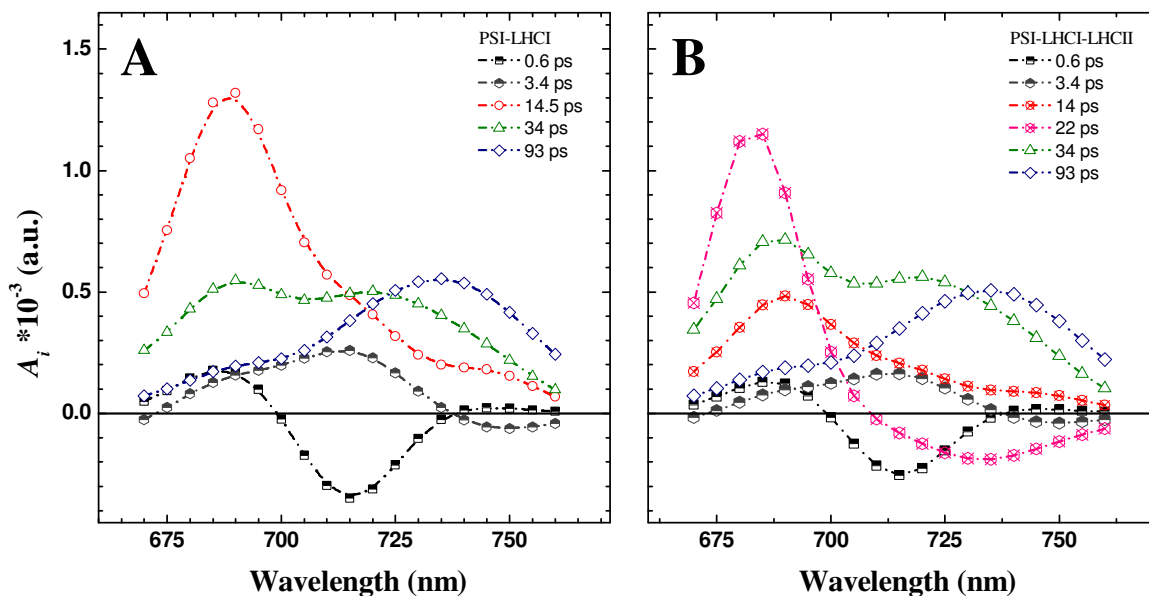

Figure S22 Complete sets of the simulated DAS in the structure-based kinetic model for PSI-LHCI (**A**) and PSI-LHCI-LHCII (**B**). The DAS that have been pooled in order to match the number of numerically retrieved spectra for PSI-LHCI (0.6 and 3.4 ps, black and grey respectively) are indicated by half-filled symbols; those that have been pooled in PSI-LHCI-LHCII simulations are indicated by either half-filled symbols (0.6 and 3.4 ps, black and grey) or crossed symbols (14 and 22 ps, red and magenta).

References

1. J. M. Beechem, E. Gratton, M. Ameloot, J. R. Knutson, L. Brand and J. R. Lakowicz, in *Topics in Fluorescence Spectroscopy. 2*, Plenum Press, New York, 1991, pp. 241-305.
2. Z. Bajzer, F. G. Prendergast, L. Brand and M. L. Johnson, in *Methods in Enzymology. 210 Numerical Computer Methods*, Academic Press, Dan Diego, 1992, pp. 200-237.
3. J. R. Alcala, E. Gratton and F. G. Prendergast, *Biophys. J.*, 1987, 51, 925-936.
4. A. Siemiarczuk, B. D. Wagner and W. R. Ware, *J. Phys. Chem.*, 1990, 94, 1661-1666.
5. G. Landl, T. Langthaler, H. W. Engl and H. F. Kauffmann, *J. Comput. Phys.*, 1991, 95, 1-28.
6. D. A. Turton, G. D. Reid and G. S. Beddard, *Anal. Chem.*, 2003, 75, 4182-4187.
7. M. Straume, S. Frasier-Cadoret and M. L. Johnson, in *Topics of Fluorescence Spectroscopy*, ed. Lakowicz, Plenum Press, New York, U.S.A., 1991, vol. Vol. 2, pp. 177-240.
8. F. James and M. Roos, *Computer Physics Communications*, 1975, 10, 343-367.
9. P. Galka, S. Santabarbara, T. T. Khuong, H. Degand, P. Morsomme, R. C. Jennings, E. J. Boekema and S. Caffarri, *Plant Cell*, 2012, 24, 2963-2978.
10. E. Wientjes, H. van Amerongen and R. Croce, *Biochim. Biophys. Acta-Bioenerg.*, 2013, 1827, 420-426.
11. E. Engelmann, G. Zucchelli, A. P. Casazza, D. Brogioli, F. M. Garlaschi and R. C. Jennings, *Biochemistry*, 2006, 45, 6947-6955.
12. R. Croce, D. Dorra, A. R. Holzwarth and R. C. Jennings, *Biochemistry*, 2000, 39, 6341-6348.
13. R. C. Jennings, G. Zucchelli, R. Croce and F. M. Garlaschi, *Biochim. Biophys. Acta-Bioenerg.*, 2003, 1557, 91-98.
14. R. C. Jennings, G. Zucchelli and S. Santabarbara, *Biochimica et biophysica acta*, 2013, 1827, 779-785.
15. P. Dainese and R. Bassi, *J.Biol.Chem.*, 1991, 266, 8136-8142.
16. R. C. Jennings, R. Bassi, F. M. Garlaschi, P. Dainese and G. Zucchelli, *Biochemistry*, 1993, 32, 3203-3210.
17. G. Zucchelli, F. M. Garlaschi and R. C. Jennings, *Biochemistry*, 1996, 35, 16247-16254.
18. A. V. Ruban, J. P. Dekker, P. Horton and R. van Grondelle, *Photochem. Photobiol.*, 1995, 61, 216-221.
19. R. C. Jennings, F. M. Garlaschi and G. Zucchelli, *Photosynth. Res.*, 1991, 27, 57-64.
20. I. Moya, M. Silvestri, O. Vallon, G. Cinque and R. Bassi, *Biochemistry*, 2001, 40, 12552-12561.
21. B. van Oort, A. van Hoek, A. V. Ruban and H. van Amerongen, *J. Phys. Chem.*, 2007, 111, 7631-7637.
22. R. Croce, G. Zucchelli, F. M. Garlaschi, R. Bassi and R. C. Jennings, *Biochemistry*, 1996, 35, 8572-8579.
23. R. Croce, T. Morosinotto, S. Castelletti, J. Breton and R. Bassi, *Biochim. Biophys. Acta-Bioenerg.*, 2002, 1556, 29-40.
24. E. Wientjes and R. Croce, *Biochem. J.*, 2011, 433, 477-485.
25. G. Zucchelli, T. Morosinotto, F. M. Garlaschi, R. Bassi and R. C. Jennings, *FEBS Lett.*, 2005, 579, 2071-2076.
26. R. C. Jennings, G. Zucchelli, E. Engelmann and F. M. Garlaschi, *Biophys. J.*, 2004, 87, 488-497.
27. E. Wientjes, I. H. van Stokkum, H. van Amerongen and R. Croce, *Biophys. J.*, 2011, 101, 745-754.