

Orientation of Chlorophyll Transition Moments in the Higher-Plant Light-Harvesting Complex CP29

Roberto Simonetto,[‡] Massimo Crimi,[‡] Dorianna Sandonà,[‡] Roberta Croce,[‡] Gianfelice Cinque,[‡] Jacques Breton,[§] and Roberto Bassi^{*‡}

Dipartimento Scientifico e Tecnologico, Università di Verona, Verona, Italy, and Département de Biologie Cellulaire et Moléculaire, CEA, Saclay, France

Received May 18, 1999; Revised Manuscript Received July 26, 1999

ABSTRACT: The Q_y transition dipole moment vectors of all eight chlorophylls in the higher-plant antenna protein CP29 were calculated by an original method on the basis of linear dichroism and absorption spectroscopy. The contribution of individual chromophores was determined from difference spectra between wild type and mutant proteins in which a single chlorophyll has been removed by mutating pigment-binding residues. Recombinant proteins were constructed by overexpressing the apoprotein in bacteria and refolding of the pigment–protein complex in vitro [Bassi, R., Croce, R., Cugini, D., and Sandonà, D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* (in press)]. The spectroscopic data are interpreted on the basis of a protein structural model obtained via the homology with the major antenna complex LHCII [Kuhlbrandt, W., Wang, D. N., and Fujiyoshi, Y. (1994) *Nature* 367, 614–621]. The results allow us to determine the orientation of six chromophores within the protein structure. The orientations of the two remaining chromophores are inferred by considering the symmetry properties of CP29 and fitting steady state absorption and linear dichroism spectra by independent chlorophyll spectral forms. As a consequence, four “mixed” sites with different chlorophyll *a* and *b* binding affinities are identified in CP29. Geometrical data and the Förster mechanism for energy transfer suggest that excitation energy equilibrates rapidly among chlorophyll “pure” sites while energy preferentially flows outward from chlorophyll “mixed” sites. The orientation of the dipole moments of two chlorophyll molecules symmetrically located at the center of the protein and parallel to the carotenoid transition vectors suggests a role in energy transfer from xanthophyll to chlorophyll.

In higher-plant chloroplasts, many pigment-binding proteins are inserted into the thylakoid membrane and organized into multisubunit complexes called photosystems. Photosystems (PS)¹ catalyze light absorption, the use of the excitation energy in transmembrane electron transport leading to the synthesis of ATP and NADPH. PS I and II are both composed of a chlorophyll *a* binding core complex and a chlorophyll *a* and *b* binding antenna. A detailed understanding of energy transfer processes in antenna and reaction centers requires knowledge of the distances between chromophores, of the mutual orientation between dipole transitions, and of the absorption and fluorescence energy levels and their distribution within the photosystems. In the case of PSII, considerable insight has been obtained from the

topological organization of subunits (*I*–*3*) and from their spectral form decomposition (*4*, *5*). Elucidation of the major antenna complex LHCII structure to 3.4 Å resolution (*6*) showed that many of the binding site environments are chemically distinct and that nearest-neighbor chlorophylls (Chl) are spaced 9–13 Å apart (center–center distance). The orientation of electronic transition dipole moments of the chromophores cannot yet be obtained from structural data, although linear dichroism (LD) reveals considerable differences in the orientations (*5*, *7*), suggesting that LD spectra could be used for transition moment vector (TMV) analysis. However, light-harvesting chlorophyll proteins (Lhc) are spectroscopically complex objects; while only one or, depending on the complex, at the most two chemically distinct Chl species are present, many spectral forms are commonly observed in the Q_y absorption region. Absorption (Abs) spectra of Lhc proteins exhibit a markedly heterogeneous broadening in the range of 630–685 nm. From optical spectroscopic measurements, 8–11 spectral forms have been identified with reasonable certainty (*8*), while the number of bound Chls per monomer is in the range of 8–14 (*6*, *9*, *10*), suggesting that the spectral heterogeneity depends on the different protein environments in which individual chromophores are located.

Insufficient progress in developing a more detailed understanding of this spectroscopic heterogeneity has been

* To whom correspondence should be addressed: Facoltà di Scienze MM.FF.NN., Università degli Studi di Verona, Strada Le Grazie, I-37134 Verona, Italy. Telephone: +39 045 8098916. Fax: +39 045 8098929. E-mail: bassi@sci.univ.it.

[‡] Università di Verona.

[§] CEA.

¹ Abbreviations: Abs, absorption; Car, carotenoid; CD, circular dichroism; Chl, chlorophyll; CP, chlorophyll–protein complex; DCCD, dicyclohexylcarbodiimide; Hepes, *N*-(2-hydroxyethyl)piperazine; DM, dodecyl maltoside; FN, fitting number; fwhm, full width at half-maximum; IEF, isoelectric focusing; LD, linear dichroism; LHCII, light-harvesting complex of PSII; Lhc, light-harvesting complex family; NPQ, non-photochemical quenching; TMV, transition moment vector; PS, photosystem; WT, wild type.

partly due to the absence of experimental techniques that are capable of selectively modifying the optical transitions. In this context, the approach of pigment–protein reconstitution is particularly interesting. Reconstitution has first been performed using apoproteins from which pigments had been detached by solvent extraction (11) and, more recently, by the overexpression of cDNA encoding single-gene products (12, 13). However, in most of these cases, the optical properties of reconstituted complexes are somehow different from those of the native ones, thus limiting their usefulness for detailed spectroscopic studies. The successful reconstitution of the minor Chl *a* and *b* antenna complexes, CP29, CP26, and CP24 using maize *Lhcb4*, -5, and -6 apoproteins overexpressed in *Escherichia coli*, was recently reported (14–16). The recombinant proteins exhibit absorption, fluorescence emission, and CD signals which closely match those of the native complex while even peculiar biochemical features such as DCCD binding (17) were reproduced (18). The recombinant CP29 system can therefore be used for the experimental analysis of the relevant parameters that have been left unknown after the structural work on LHCII by Kuhlbrandt and co-workers. In fact, sequence analysis showed a high degree of homology between LHCII and CP29 (19), while the Q_y transition spectral forms, resolved in the spectroscopy of LHCII and CP29 proteins, were essentially identical (5, 20). The following major points are investigated here: (a) identification of the energy levels of the individual Chl molecules, i.e., which of them are Chl *a* and which are Chl *b* and at which wavelength they absorb; and (b) determination of the orientation of the transition moments of individual Chl molecules. In an earlier stage of this work, the first problem has been approached by introducing single-point mutations of Chl binding residues into the polypeptide sequence to construct recombinant mutant proteins missing each one of the chromophores. Biochemical analysis showed that most of the mutants bind seven Chl molecules per polypeptide versus eight in the wild type (WT) protein.

In this study, we address the question of the orientation of the Q_y transition of the individual Chl molecules in CP29. To this aim, single-point mutants were analyzed by low-temperature linear dichroism and absorption spectroscopy. These data were used for screening the different possible orientations for each chromophore obtained from a structural model of the CP29 protein obtained via its homology with the LHCII structure (6). The reconstruction of both Abs and LD spectra of the WT protein in the Q_y region allowed the determination of the TMV of Chls that could not be accessed experimentally by mutation analysis and supports the previous suggestion that in CP29 some binding sites have mixed Chl *a* and *b* occupancy (21).

MATERIALS AND METHODS

DNA Constructions. Plasmids were constructed using standard molecular cloning procedures (22). Bacterial hosts were the *E. coli* TG1 strain (23) and the *E. coli* SG13009 strain (24). Mutations were introduced according to the method of Yukenberg et al. (25). The sequence was determined by the dideoxy method (26) by an automated apparatus (Applied Biosystems model 377 sequencer).

Isolation of the Overexpressed CP29 Apoprotein from Bacteria. CP29 was isolated from the SG13009 strain

transformed with either of the two CP29 constructs following a protocol previously described (12, 27).

Reconstitution of CP29. These procedures were performed as described by Giuffra et al. (14).

Purification of Reconstituted CP29 and LHCII. This was performed by ion-exchange chromatography (14). For the determination of pigment to protein stoichiometry, it was necessary to obtain a fully purified protein which did not contain any residual contamination by bacterial proteins. The reconstituted CP29 was thus purified by preparative IEF (28) followed by ultracentrifugation in a glycerol gradient [15 to 40%, including 0.06% DM and 10 mM Hepes (pH 7.6); the run lasted 12 h at 60 000 rpm in a SW60 Beckman rotor] to eliminate ampholytes.

Protein and Pigment Concentrations. The concentration of the CP29 apoprotein purified from *E. coli* inclusion bodies was determined by the bicinchoninic acid assay (29). For the stoichiometric (pigment/protein ratio) determination, the protein concentration was determined by the ninhydrine method (30). The chlorophyll concentration was determined by the method of Porra et al. (31). HPLC analysis was carried out as described by Gilmore and Yamamoto (32).

Spectroscopy. Absorption spectra were obtained using a SLM-Aminco DW-2000 spectrophotometer at room temperature. Samples were in 10 mM Hepes (pH 7.6), 0.06% DM, and 20% glycerol. The chlorophyll concentration was about 10 $\mu\text{g}/\text{mL}$ for absorption measurements and 0.01 $\mu\text{g}/\text{mL}$ for fluorescence measurements. The 100 K absorption and linear dichroism spectra of the recombinant LHCII proteins have been measured according to the method of Haworth et al. (7) using samples oriented by the polyacrylamide squeezing technique as previously described (5, 33).

RESULTS

CP29 Structural Model. For the construction of a structural model of CP29, the information that is available consists of the α -carbon coordinates of the membrane-spanning helices and the raw structure of 12 tetrapyrrole rings and two xanthophylls in LHCII kindly provided by W. Kuhlbrandt (6). By using the computer program MaxSprout (34), which is available from the European Molecular Biology Laboratory, we proceeded to reconstruct the spatial organization of amino acid side-chains. In principle, the exact positions of the side chains of each residue are not indispensable for the interpretation of LD, but they are useful in evaluating the quality of the model. In fact, the residues found to be involved in Chl binding by site-directed mutagenesis are expected to have a spatial location consistent with coordination to tetrapyrroles. The CP29 structural model was therefore obtained by substitution of the CP29 amino acid residues in the LHCII structure on the basis of the alignment in Figure 1, which also shows the conservation of the eight Chl binding residues (or ionic pairs) identified in LHCII (6). Mutagenesis of each of these residues in CP29 confirmed their role in Chl coordination (21). Since CP29 binds eight chlorophylls per polypeptide, the four additional tetrapyrroles, which did not show a suitable coordinating residue, were omitted from the model. The CP29 structural model is represented in Figure 2.

Absorption and LD Spectroscopy. Figure 3 shows the 100 K Abs and LD spectra of CP29 WT and point mutants in

Chl Site	A4	A5	A1	
LHCII - helix B	PETFSK NRE LEVIH SR WAMLGALGCVFPPELLSR			
CP29	VFGLQRFR EC ELIH GR WAMLATLGLALSVEWLTG			
	. . . * * * . . . * * * . . . * * .			
42% identity				
Chl Site	B6	B5	B5	
LHCII - helix C	SILA-IWATQVVL M GAVEGYRI			
CP29	SISTLIW-I EV LVI G YIEFQ R N			
	** . ** . . . * . . . * * *			
38% identity				
Chl Site	A1	A2	A4	A3
LHCII - helix A	PEAF A ELK V EL K NGRLAMF S MF G FFV Q A I			
CP29	PEK K ER L Q L A E IK H AR L AM V AF L GF A V Q A A			
	** . * . . * . . . * * * . . . * * * * * *			
47% identity				
Chl Site	B3			
LHCII - helix D	PLENLAD H LA			
CP29	PLNNW A THLS			
	** . * * * . . .			
50% identity				

FIGURE 1: Alignment of the four α -helix regions of CP29 (sequence from *Zea mays*) with the homologous LHCII sequences as resolved in the electron crystallographic atomic model (6): (*) identical residue and (•) conserved residue. Bold letters represent residues shown to be involved in chlorophyll coordination (6, 21).

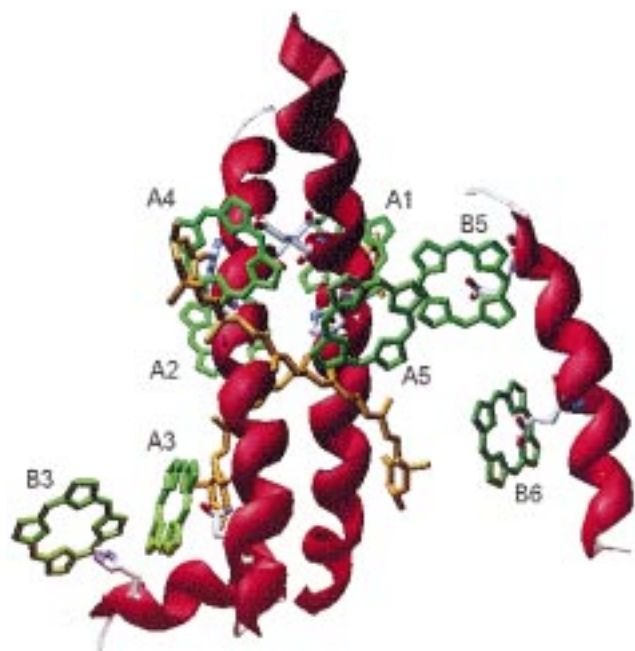


FIGURE 2: Structural model of the CP29 protein.

which a single pigment-binding residue has been changed to a nonbinding one. We could not obtain a mutant protein missing the chromophore in site A1 (see Figure 2). The pigment in this site has been shown to be coordinated by the Arg116–Glu213 ionic pair. Mutation of either one or both of the residues forming the ionic pair prevented *in vitro* reconstitution (21). Moreover, the mutant protein Q230L, although exhibiting difference absorption spectra and biochemical properties consistent with the loss of a single Chl (in site A3), yielded a LD signal that was unsuitable for further analysis. From these data, we proceeded to the identification of the contribution of each Chl to the total Q_y transition in both Abs and LD data.

Difference absorption spectra between WT CP29 and the individual mutants were obtained following normalization of the total integrated area under the Q_y absorption band using the measured Chl *a* and Chl *b* protein stoichiometries and

assuming a relative Chl *b*/Chl *a* extinction ratio of 0.7 (20). Over the wavelength range that was analyzed, all the difference spectra have the expected positive amplitude except for two cases in which the small negative contribution at around 680 nm is likely due to a baseline drift. In most cases, WT minus mutant difference spectra could be fitted with one major and one or two minor Gaussian components peaking at the frequencies previously reported for Lhc chlorophyll absorption (data not shown). Such a result suggests that the removal of one pigment from the complex did not introduce major changes in the spectral properties of the remaining chromophores. This is in agreement with previous work in which the effect of the protein binding site was shown to be the major determinant of pigment absorption (20, 21). A representative example of a WT minus mutant Abs spectrum is shown in Figure 4 (solid line) for the mutant H216F.

Normalization of LD Spectra. To normalize the LD spectra of each mutant with respect to the WT spectrum, it is not possible to use the same method that was used for the Abs spectra since the intensity of each component of the LD spectra depends on the angle of the chromophore with the normal to the plane in which the protein is oriented. In other words, the amplitude is not solely related to the chromophore concentration. Moreover, the evaluation of any normalization factor is complicated by the fact that the LD signal can be either negative or positive and because the extent of orientation of the proteins in the gel may vary from sample to sample.

To overcome these problems, we developed a software code that accomplishes normalization of the LD spectra from the information derived from the WT minus mutant difference Abs spectra. In fact, the LD signal of a single chromophore transition is expected to have the same shape as its Abs signal multiplied by a constant whose sign and amplitude depend on the orientation. On this basis, we can expect that a correct normalization will lead to a WT minus mutant LD spectrum with amplitudes proportional at each point to the normalized difference Abs spectrum between the WT and the same mutant. Exploiting this property for the Q_y transition of the Chl affected by the mutation, the algorithm proceeds to the exploration of possible normalization factors of mutant protein LD. These spectra are subtracted from the WT LD spectrum, thus yielding a family of difference spectra. For the evaluation of their relative quality, the procedure compares each of them to the WT minus mutant Abs spectrum by searching for the minimal difference in a wavelength window selected to include the main Q_y transition of the missing Chl. LD spectra normalized by this factor are used to obtain WT minus mutant difference spectra. For all the difference LD spectra of the individual mutants being calculated using the same WT LD spectrum, these are thus normalized relative to each other for any difference in the extent of orientation of the complex in the gel.

All the resulting LD difference spectra are positive or negative over the entire region in which the computation has been performed, supporting the homogeneity of the spectral forms identified in this manner. An example of the results obtained with this procedure is shown in Figure 4 (dotted line) for the H216F mutant. Table 1 summarizes the results of this procedure for the six mutant proteins that were

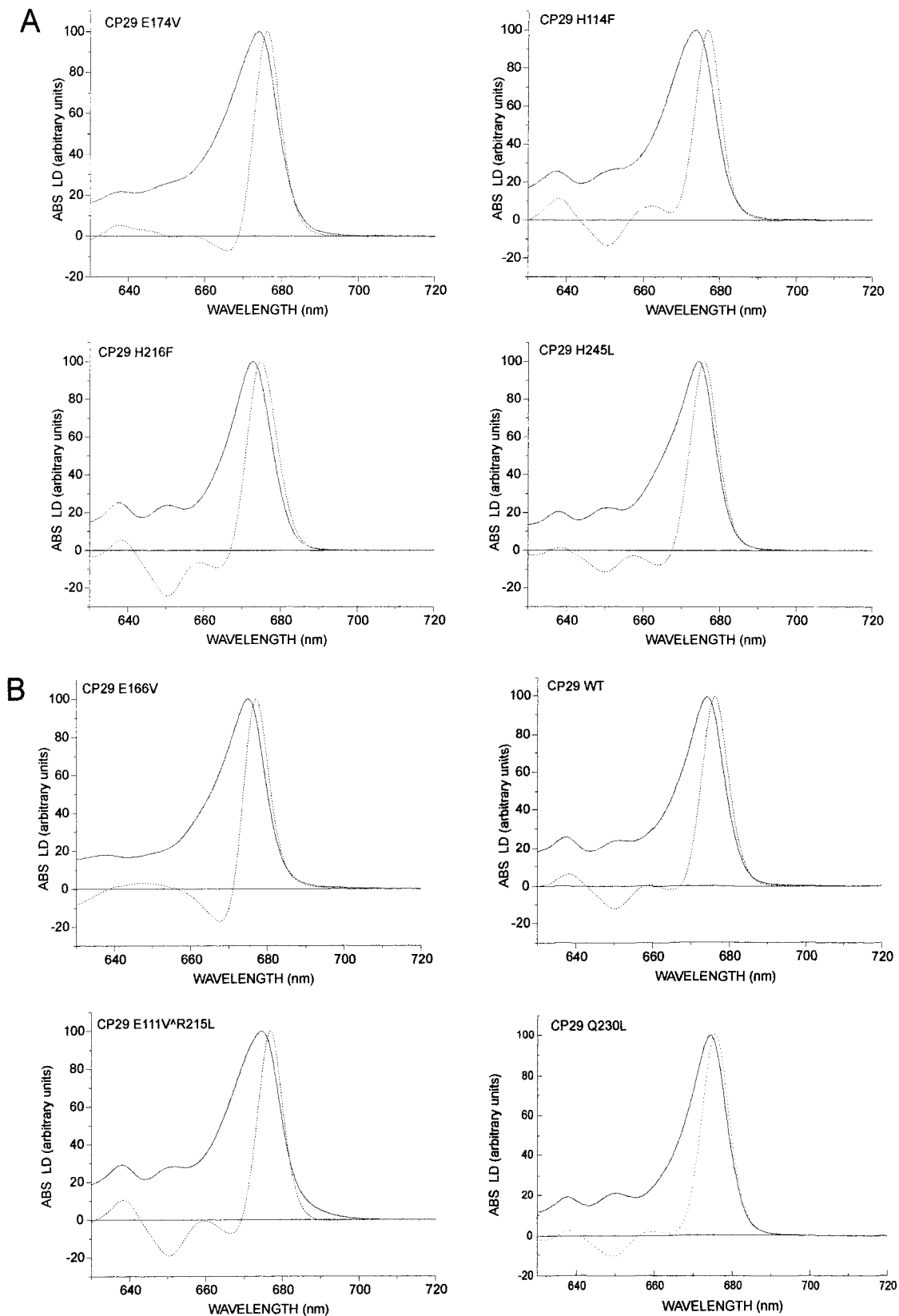


FIGURE 3: Absorption and linear dichroism spectra at 100 K of the seven mutants and wild type recombinant CP29: (—) absorption spectrum and (···) LD spectrum.

analyzed. In one case, the Q230L mutant, we could not obtain a reliable normalization due to the low amplitude of the LD signal. Finally, it was necessary to normalize the LD spectra with respect to the Abs spectra. Since the spectra within these two groups have been normalized with respect to the WT spectra, the factor between absorbance and LD amplitudes

should have the same value for all the calculated WT minus mutant difference spectra. This normalization factor, represented by FN, is related to the mosaicity of the WT protein within the acrylamide gel in which the protein was included for orientation by the compression technique as discussed in ref 33.

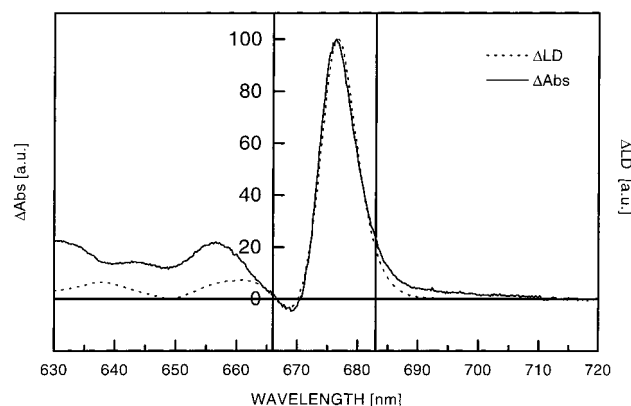


FIGURE 4: Wild type minus H216F mutant absorption and linear dichroism difference spectra after normalization. The wavelength window for the best fit software search is indicated by vertical bars (see the text).

Table 1: Best Fit Results between Absorption and Linear Dichroism Wild Type Minus Mutant Difference Spectra^a

Chl site	CP29 mutant	$\lambda_{\min}-\lambda_{\max}$ (nm)	A_{Abs} (arbitrary units)	A_{LD} (arbitrary units)
A2	H216F	666–683	1.5	379
A4	E111V/R218L	666–677	1.85	155
A5	H114F	668–691	1.31	240
B5	E174V	644–657	0.29	–104
B3	H245L	630–640	0.42	68
B6	E166V	665–679	1.94	295

^a Mutant proteins and the corresponding wavelength window used for the best fit of WT minus mutant linear dichroism to WT minus mutant absorption difference spectra are shown. The last two columns report the area under absorption and linear dichroism difference spectra after normalization.

The relationship between the intensity of the linear dichroism and the relative absorption signal, in the form given by Breton et al. (33), is therefore modified as follows:

$$\text{LD} \equiv A_{\parallel} - A_{\perp} = \frac{3}{2}A \times \text{FN}(1 - 3 \cos^2 \varphi) \quad (1)$$

where φ is the angle between the normal to the complex plane and the chromophore transition moment vector. Clearly, the LD provides information only about the angle between the TMV and the normal to the plane in which the protein is oriented. Since during the preparation of the samples for LD measurement membrane proteins aggregate via their hydrophobic side surfaces, the orientation that is obtained can be considered to be the same as that determined by electron microscopy in two-dimensional layers (6, 35). The LD data analysis was therefore performed by using as a normal to the membrane plane the 2-fold symmetry axis defined in the LHCII crystallographic structure. The dichroic and the absorption signal amplitudes employed in such calculations are the intensities integrated in the region of the spectra involved in the transition, which is the same as that used for the procedure of normalization.

Due to the lack of information about the ring V and phytol chain esterification position (Figure 5) in the tetrapyrrole rings determined in the LHCII structure (6), which would define the TMV, a free search in the space of possible chlorophyll orientations must discriminate among 8^8 cases for the CP29 protein. It is therefore important to reduce the number of acceptable solutions that will be evaluated.

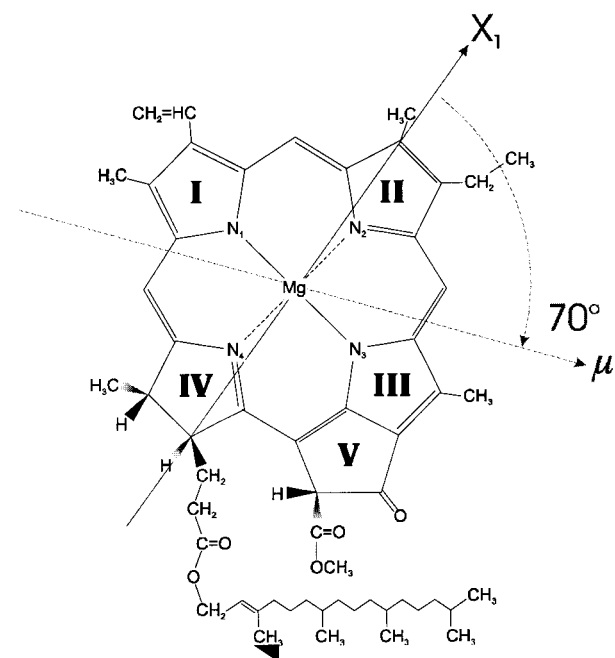


FIGURE 5: Orientation of the Q_y transition moment vector (μ) and X_1 chlorophyll axis definition.

Toward this aim, it is useful to consider the behavior of the phytol chain in antenna systems whose structure is known in greater detail; this is the case of the bacterial LHIII complex whose crystallographic structure has been determined at high resolution (36, 37). The organization of two (bacterio)-chlorophylls within the membrane is similar to that determined in LHCII. Here the tetrapyrrole plane is perpendicular to the membrane plane, while the pigment is oriented with the phytol chains directed toward the internal (more hydrophobic) region of the membrane. This implies that phytols are esterified to the pyrrole substituents directed toward the inner (more hydrophobic) part of the membrane layer. The opposite orientation with chains aiming toward the outside of the membrane is extremely unlikely, since it would imply an abrupt bending of the chain for re-entry into the membrane. For the above reasons, we have excluded the orientations of the TMV corresponding to phytols oriented toward the membrane surface. Since a clear-cut distinction between external and internal sites is possible, this procedure leaves only four possible positions of the phytol esterification site for each Chl in CP29, i.e., a total of 4^8 possible orientations.

A geometrical system was constructed for the calculation of the possible TMV for each Chl. In this system, the coordinates of three atoms for each tetrapyrrole are introduced to define the geometrical plane to which the TMV belongs (see Figure 5). Two of these atoms are chosen to be carbon 7 in pyrrole IV and carbon 3 of pyrrole II, which define the X_1 axis. Our knowledge of the orientation of the electronic transition with respect to the molecule-fixed coordinate system is based mainly on the results of LD studies with molecular Chl solutions oriented in different systems. The main Q_y transition of Chl *a* has its maximum at 670 nm. It was shown that its TMV is lying in the plane of the tetrapyrrole ring and is defined with an angle measured in degrees clockwise with respect to the X_1 axis (Figure 5). The Q_y orientation has been shown experimentally to form an angle with respect to the X_1 axis of 70° for Chl *a* and 61°

Table 2: Solutions for the Azimuth Angles and Corresponding Fitting Numbers^a

Chl	φ (deg)	FN	Chl	φ (deg)	FN	Chl	φ (deg)	FN
A2	79	190	A4	90	56	A5	71	183
	<u>75</u>	<u>212</u>		<u>60</u>	<u>221</u>		84	<u>126</u>
	35	-168		34	-53		31	-103
	31	-138		13	-30		38	-141
B5	74	-309	B3	47	-270	B6	81	110
	79	-271		56	2772		68	<u>177</u>
	29	187		54	-8510		23	-65
	<u>32</u>	<u>203</u>		<u>65</u>	<u>229</u>		10	-53

^a Data were obtained from the structural model of CP29 on the basis of four different positions of phytol chain connection to the tetrapyrrole ring. Underlined values are recurring solutions for the Q_y transitions of the Chl molecules that were analyzed; within the error, these solutions satisfy the requirement of a common FN value (see the text).

for Chl *b* (38, 39). Two other values of 110° and 119° for Chl *a* and Chl *b*, respectively, must also be considered because of the two possible faces of a Chl.

Fitting the Experimental LD Data into the CP29 Molecular Model. For each of the six Chl molecules whose Abs and LD spectra have been considered here, the CP29 structural model provides four anchoring points for the phytol chain. This in turn gives four possibilities for the angle between the TMV of these Chl molecules and the normal to the complex plane, when taking into account the α angles of 70° and 61° in the internal coordinates of the TMV of Chl *a* and Chl *b*, respectively. For each of these angles, a theoretical LD/3A value can be calculated. Upon comparison with the corresponding LD/3A experimental value for that particular Chl, a FN value is obtained. The FN values calculated for the six Chl molecules analyzed by LD spectroscopy are listed in Table 2. Within this set of FN values, only those in common with all the chromophores that were investigated can be considered as possible solutions. We therefore searched for a recurrent value of FN within each group of four solutions. Such an FN value was found (ranging between 177 and 229) with two possible solutions for both chromophores in sites A2 and B5 (Table 2). No common values of FN were obtained upon varying in steps of 10° the value of the α angle (between 60° and 100° for Chl *a* and between 51° and 91° for Chl *b*). Again, no common values of FN could be found upon using in the computation the alternative angles considered by Fragata et al. (38) for the Q_y transition (110° and 119°). The final value of 204 for FN has been estimated to give the minimal root-mean-square deviation within the group of selected results. The corresponding standard deviation for the TMV angles has a value of 1.7°.

Inferring Chl A2 Orientation from CP29 Internal Symmetry. For the A2 site, the multiplicity of TMV solutions obtained by LD and Abs analysis of mutant H216F can be disentangled by considering the symmetry properties of the whole protein. Thus, with respect to the membrane plane normal of CP29 (see Figure 2), the X-shaped structure constituted by two α -helices (A and B) has a local 2-fold symmetry which, together with the high degree of homology in the first 24 amino acid residues of both helices (6, 40), is reflected in a symmetrical positioning of A1 to A4 and A2 to A5 Chl planes.

Since the TMV of Chl A5 has been determined from the differential spectroscopy of CP29 WT versus the H114F

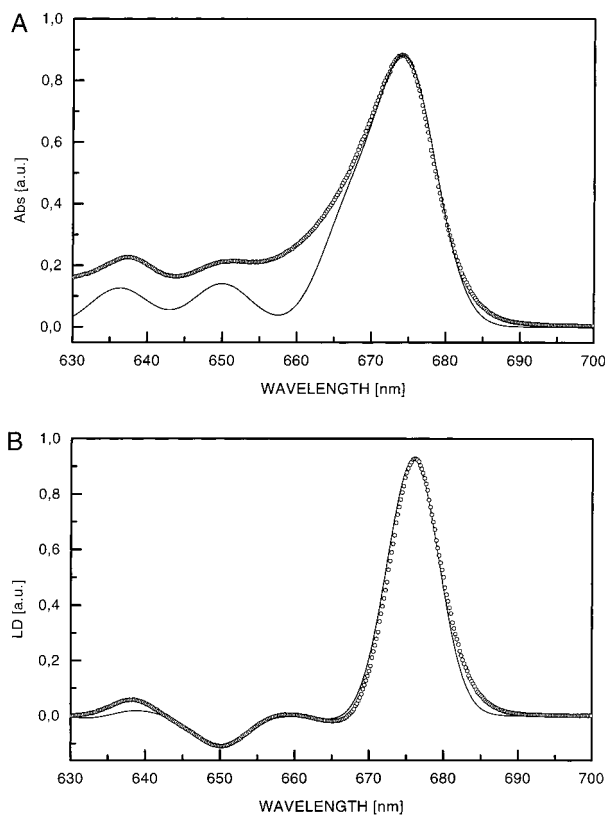


FIGURE 6: Reconstruction of absorption and linear dichroism spectra for CP29 WT at 100 K with eight independent Gaussians associated with Chl pigments. Equation 1 in the text was used for relating absorption to linear dichroism for each sub-band: (—) reconstructed spectra and (○) experimental spectra.

mutant, the orientation of Chl A2 TMV can be determined by its symmetry with A5 with respect to the membrane normal axis and by its similarity to the relative orientation of A4 to A5 TMVs. Such an orientation coincides with the second choice, i.e., φ being equal to 75° in Table 2, of the available solutions for Chl A2 in the geometrical system.

Reconstruction of Experimental LD and Absorption Spectra. The results described above greatly reduce the number of possible orientations for the remaining two chromophores in sites A1 and A3 whose LD properties could not be determined in this study because of the low quality of the spectrum in the case of site A3 and because of the instability of the mutant protein in the case of site A1. In a simple approach, we have attempted to reconstruct the 100 K steady state LD and Abs spectra of WT CP29 by summation of independent Gaussian components with a unique fwhm (about 8–9 nm) and transition wavelength as experimentally determined from difference absorption spectra (Figure 6A,B). In associating one spectral form per pigment, the LD spectral amplitude of each component is weighted by the orientation of the corresponding Chl TMV according to eq 1. The symmetric spectral form at 100 K follows from the studies of sub-band absorption decomposition (41), and the extinction ratio of 0.7 for the Chl *b* to Chl *a* sub-band area in the Q_y region has been used.

From differential absorption, it was possible to locate the transition of Chl *a* in site A3 at 667 nm. A mutant lacking the chromophore in site A1 was not available. However, by subtraction of all the difference absorption spectra (WT minus mutant) from the CP29 WT spectrum, a sharp major

Table 3: Chl Parameters from Linear Dichroism and Absorption Spectral Reconstruction of WT CP29^a

Chl site	$f_{\text{Chl}-a}$ (%)	$\lambda_{\text{Chl}-a}$ (nm)	$1 - 3\cos^2\varphi$	$f_{\text{Chl}-b}$ (%)	$\lambda_{\text{Chl}-b}$ (nm)	$1 - 3\cos^2\varphi$
A1	100	669	0.99	—	—	—
A2	100	676.6	0.80	—	—	—
A3	70	667	-1.39	30	636	-1.56
A4	100	674.2	0.26	—	—	—
A5	100	674.6	0.65	—	—	—
B5	60	674	-1.17	40	650	-1.35
B3	30	673.2	0.69	70	636.4	0.44
B6	40	674.4	0.19	60	650	0.57

^a Chlorophyll *a* and *b* occupancy fractions in Chl sites together with their main associated Q_y transition and orientation factor $1 - 3\cos^2\varphi$ appearing in eq 1. The sign of the values in the last column gives the sign of the corresponding linear dichroism sub-band.

component peaking at 669 nm was obtained. Therefore, the major absorption contribution associated with the A1 chromophore was tentatively located at this wavelength. Fitting the LD spectrum required discrimination among four possible orientations of transition moments for each of the chromophores A3 and A1. The only satisfying solution was found using the A1 and A3 orientations reported in Table 3, the other possibilities yielding reconstructed spectra that were strikingly different from the experimental ones. Notice that the TMV of Chl A1 is constrained by the same 2-fold symmetry of Chl A1 and A2 on helix A and Chl A4 and A5 on the helix B subcomplex; in fact, the A1 TMV is set in agreement with the symmetric Chl A4 TMV. The orientation of Chl *a* in site A3 was imposed by the negative contribution that is necessary at the transition wavelength of 667 nm in the LD spectrum. Although the fitting at this stage was reasonably good, the ratio of the main Chl *a* peak amplitude with respect to the Chl *b* components between 640 and 650 nm in the reconstructed LD was clearly lower than that from the experimental spectrum. The final spectra of panels A and B of Figure 6 were obtained by the necessary introduction of a mixed occupancy by Chl *a* and Chl *b* of sites A3, B3, B5, and B6 (Table 3) as determined in previous work (21). Notice that the difference absorption WT minus mutant spectra of these sites already showed simultaneously two (or more) major components below and above 660 nm. On the other hand, the mutants of sites A2, A4, and A5 give a single major peak above 660 nm (21). Again, the LD reconstruction discriminates between the two possible TMVs in the B5 site, solutions of the geometrical system, and favors the most negative contribution (i.e., φ equal to 32° in Table 2). The reconstruction reproduces all the features of LD and Abs spectra with respect to both the transition wavelengths and relative amplitudes. Since we did not make any attempt to introduce vibrational transition, the absorption offset observed in the Q_y region between 630 and 660 nm is related to the missing higher-energy states of Chl *a* Gaussian components. Table 4 reports the orientation of transition moments for the chromophores in each of the eight porphyrin sites of CP29 as a result of direct LD measurements (sites A4, A5, B3, and B6), or accompanied by symmetry reasons (site A2), and fitting of steady state LD and Abs (sites A1, A3, and B5).

DISCUSSION

The question of determining the orientation of chromophores in Lhc proteins involves two essential require-

Table 4: Transition Moment Vectors of Pigments in the CP29 Complex^a

pigment site	pigment type	l	m	n
A1	chl <i>a</i>	-0.88	-0.48	0.07
A2	chl <i>a</i>	-0.88	0.39	0.25
A3	chl <i>a</i>	-0.27	0.36	0.89
A3	chl <i>b</i>	-0.32	0.22	0.92
A4	chl <i>a</i>	-0.65	-0.57	0.45
A5	chl <i>a</i>	0.88	-0.34	0.34
B5	chl <i>a</i>	0.17	0.50	-0.85
B5	chl <i>b</i>	0.02	0.47	-0.89
B3	chl <i>a</i>	-0.94	0.13	0.32
B3	chl <i>b</i>	-0.90	0.03	0.43
B6	chl <i>a</i>	0.64	-0.56	-0.52
B6	chl <i>b</i>	0.70	-0.61	-0.38
Lut1	Car	-0.74	0.39	0.55
Lut2	Car	0.77	-0.40	-0.51

^a Transition moments are expressed as versor components (l , m , n) with respect to the Cartesian axis of the modified LHCII PDB structure file. The z -axis is the normal to the membrane plane; i.e., $n = \cos\varphi$ in eq 1 of the text.

ments: (i) an experimental system in which individual chromophores can be modified and therefore analyzed and (ii) a structural model from which the azimuthal angle φ in eq 1 can be determined. At the present stage of knowledge of Lhc proteins, structural data are only available for LHCII (6). The use of LHCII as an experimental model is, however, hampered for a number of reasons. First, LHCII is a trimer, and the interactions between chromophores belonging to neighboring monomers are strong and modify the spectral properties of the system (42–44). Second, LHCII binds 12 Chl molecules per polypeptide and three xanthophylls, while only eight Chl binding sites and two xanthophyll binding sites have been identified so far in the protein structure. Experimentally, the binding sites of four Chls and one xanthophyll cannot be selectively modified in this protein at present. Finally, LHCII, as purified from plants, is heterogeneous, being a mixture of many (up to 18) gene products. It is therefore difficult to determine whether the recombinant LHCII system suitably reproduces the properties of native LHCII.

In these respects, CP29 is a better system than LHCII for the following reasons: (a) it is a monomeric protein, (b) it binds eight Chl and two xanthophyll molecules whose ligands can be identified and experimentally modified, and (c) it is a single-gene product, and therefore, we have been possible to check that recombinant and native CP29 are identical in their spectroscopic and biochemical properties. The need for a structural model of CP29 was approached by constructing a model using its homology with LHCII. The structure that was obtained (Figure 2) is a model in which the coordinates of the tetrapyrrole rings are those experimentally obtained from the two-dimensional crystals of LHCII; given the high degree of homology that exists within this protein family (19) which leads to closely similar spectral properties (5), it is reasonable to use such structural information for the calculation of transition moment vectors in CP29. This approach received a posteriori support from the successful reconstruction of both Abs and LD of WT experimental spectra using data for Chls and coordinates derived from the LHCII structure.

Linear dichroism has been long used for obtaining information about the orientation of chromophores in pig-

ment–protein complexes. When structural data at atomic resolution are available, a quantitative analysis could be achieved. This is not the case with Lhc proteins except for LHCII, whose detailed analysis of LD is difficult due to the presence of a large number of spectrally overlapping chromophores and to the lack of an experimental system allowing selective manipulation of individual pigments. This limitation has recently been overcome in the case of the homologous protein CP29 for which an *in vitro* reconstitution system has been developed, yielding recombinant proteins spectroscopically indistinguishable from the native complex (14, 20). By this approach, the spectral contribution of each component chromophore could be isolated using direct mutagenesis of the pigment-coordinating amino acid residues (21). This study reports on the results of the first attempt to correlate a structural model of an Lhc protein with LD data obtained from mutant proteins in which a single Chl chromophore has been removed at the individual binding sites.

The absorption and LD results obtained from six CP29 mutant proteins are consistent with essentially one combination of TMV orientations of four Chl *a* and two Chl *b* chromophores. This allows us to propose a position for the anchoring of the phytol chain to the tetrapyrrole ring of each of the six chlorophylls (blue dot in Figure 7). In two cases (site A2 and site B6), the LD results permit only the assignment of the pyrrol ring binding the phytol chain but not discrimination between C7 and C8.

It is interesting to note that an allowed solution was found only when using in the calculation α angle values of 70° and 61° defining the Q_y transition with respect to the chlorophyll-fixed coordinate system, while no solutions were obtained by using the complementary angles (110° and 119°), thus supporting the first set of values among the two provided by the previous studies using chlorophylls oriented in lipid bilayers (38, 39). Similarly, no allowed solutions were obtained when different values of the angle ranging between 60° and 100° for Chl *a* and 51° and 91° for Chl *b* were used.

The symmetry properties of the CP29 complex in the two α -helices A and B and their homologous binding sites for the couples A1–A2 and A4–A5 have been invoked for choosing between the two TMV possibilities for A2. To some extent, these same properties helped in orienting A1 TMV in space.

However, the main improvement in orienting the TMV of CP29 Chls was obtained by the reconstruction of the WT optical spectrum in the Q_y region. In the frame of independent pigments, the overlap of single Chl *a* and Chl *b* absorption nicely accounts for the Abs spectrum of the whole complex (Figure 6A). The underestimated Abs at higher energies is related to the fitting with Gaussian bands, which are well suited for broadened electronic transitions but neglect vibrational states. The dichroic signal of each pigment is directly proportional to the absorption sub-band throughout eq 1, and the excellent LD fitting is shown in Figure 6B. The successful reconstruction of Abs and LD by the same set of eight sub-bands, one per each binding site, confirms the hypothesis of the noninteracting pigment underlying the differential spectroscopy of single-point mutants; in this sense, the absence of a significant coupling between Chls in CP29 Abs and LD spectra has already been discussed by Giuffra et al. (20). The transition wavelength and orientation

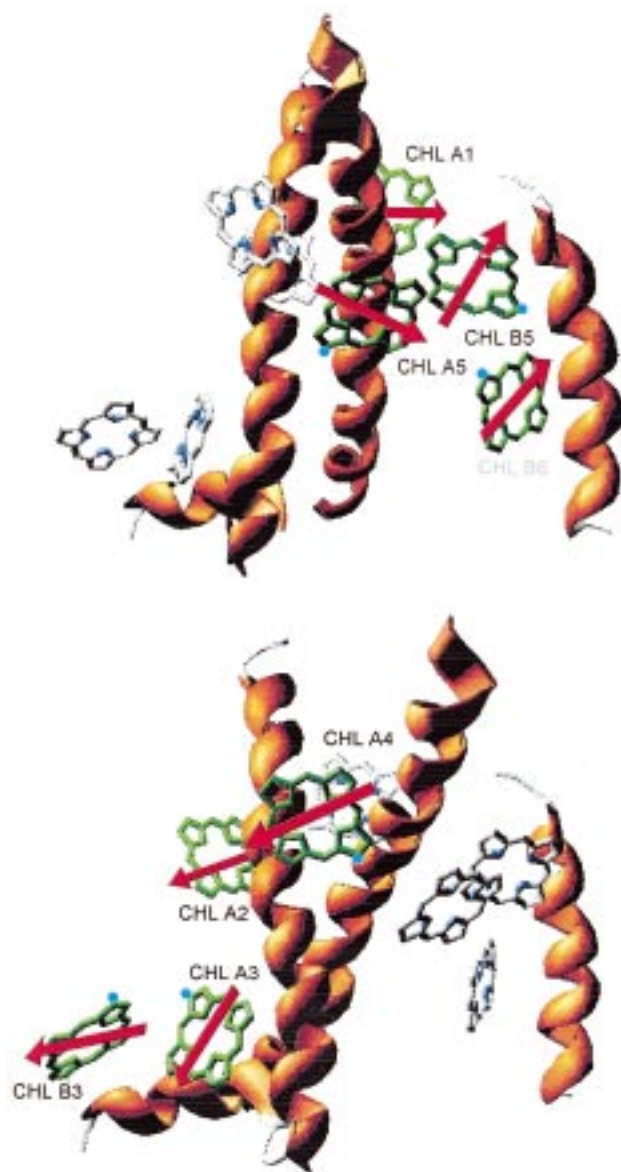


FIGURE 7: Representation of transition moment vectors (red arrows) of all the chlorophyll molecules in the structural model of CP29. For the sake of clarity, the complex is shown in two different views rotated about 20° around the vertical axis. The blue dots on tetrapyrroles represent the proposed binding site of the phytol chain (Chl A1 and A5 phytol binding sites hidden by α -helices).

parameters that were used for the spectral reconstruction are reported in Table 3. The assignment of the optical transition to each Chl binding site in antenna systems is still an open question which was widely studied by optical spectroscopy and Gaussian decomposition either in the case of CP29 (20) or more generally for the Lhc family in PSII (5, 16). The single-site mutational approach coupled to the spectroscopic analysis is a powerful tool in pinpointing the contribution of missing Chl in the differential spectra (WT vs mutants, see panels A and B of Figure 3). In this labeling, it is interesting that some sites appear to be partially occupied by Chl *a* and *b*; this is a necessary requisite for reproducing the WT Abs and LD experimental spectra and satisfying the total pigment stoichiometry of six Chls *a* and two Chls *b* per protein. While mutants lacking a Chl in A2, A4, and A5 give single Abs peaks in the window between 670 and 680 nm (e.g., Figure 4), mutations in sites B5 and B6 clearly

Table 5: Distances between Chlorophylls in the CP29 Complex^a

$d_{\text{Chl-Chl}}$ (Å)	A2	A3	A4	B5	A6	B3	B6
A1	12.3	26.2	16.2	19.4	16.4	31.4	20.3
A2		19.2	17.4	24	24.9	21.4	25.5
A3			19.6	21.3	27.5	9.2	27
A4				12.7	17.5	25.2	26.9
A5					8.9	29.9	18.8
B5						35.8	14.6
B3							35

^a Mg atom distances for the chlorophylls present in CP29 as modeled by its homology to the LHCII experimental structure.

present two sub-bands at both 650 and \sim 675 nm. The differential spectrum of the Chl B3 mutant exhibits more than two peaks below and above 660 nm. In brief, the binding sites for Chl A2, A4, and A5 seem to be selective for Chl *a*; the same homology of helices A and B of CP29, specifically in the residues binding these Chls, strongly suggests that site A1 is also selective for Chl *a*. Data for mutants of B5, B6, and B3 result in the most unrestrictive hypothesis that these sites have an affinity for both Chl *a* and *b*, and the fraction of mixed population has been quantified by LD reconstitution in agreement with previous biochemical analysis (21).

A graphical representation of the Chl TMV in the CP29 antenna protein is shown in Figure 7. Table 5 summarizes the center–center distances between the eight chlorophylls, while Table 4 shows the orientation of the Q_y transitions and the two xanthophyll TMVs obtained from the direction of the polyene chain. Since the TMV orientations and distances between chromophores in the CP29 complex are available, a few words about the energy transfer and excitation pathway among the ensemble of Chls are in order. According to the Förster mechanism, the overlap integral between the donor fluorescence spectrum and acceptor absorbance is also a part of the calculation and describes the dependence of the transfer rate on the temperature; thus, the following comments refer exclusively to the geometrical factor accounting for the relative dipole–dipole orientation and the scaling like the sixth power of the distances between the pigments. In this sense, the fastest excitation transfer takes place between Chl A1–A2 and A4–A5, while their cross-interaction should be slower but certainly effective. Notice that this assembly of easily equilibrated pigments is exclusively composed of Chl *a* in nonmixed states, and if a kinematic model of energy transfer is assumed, it is surrounded by all the other mixed sites. Chls in such mixed sites have a much slower rate of transfer of excitation, and the transfer is almost exclusively to the two closest neighbors. Clearly, the geometrical factor and relative distance provide no information about the manner of excitation flow; this means that when all sites are filled by identical Chls, which is possible only for Chl *a*, the process must be bidirectional. In consideration of the Stokes shift accompanying fluorescent emission and of the large energy difference between Chl *a* and *b* Q_y transition, the energy transfer from mixed sites with Chl *b* occupancy to the nearest Chls *a* is favored. The exception is represented by the mixed site B5, where both subpopulations of Chls *a* and *b* are strongly coupled to Chl *a* in A5 by a transfer rate that is comparable to that of the faster Chl A–A site interaction. In recent experimental and theoretical studies on the energy transfer in monomeric LHCII (45) and CP29 (46), the simultaneous presence of a

fast and slow equilibration process among lower-energy Q_y transitions, or equivalently Chls *a*, was already envisaged; this is in agreement with our modeling of a Chl *a* assembly with a high rate of transfer between A1–A2 and A4–A5 and a slower intercrossing equilibration. Notice that this assembly of Chls *a* is conserved in the structure of CP29 and LHCII; in fact, experimental evidence of both rapid and slow equilibration between some Chl *a* molecules in LHCII at low temperatures was found by Visser et al. (47). As for the Chl *b* excitation in transient absorption at 77 K, the largest energy transfer constant was between the blue-most Chl *b* and the red-most Chl *a*, while the opposite uphill process (red-most Chl *b* to blue-most Chl *a*) was found to be 1 order of magnitude slower (46). In our kinetic model, the B5 site Chl *b* population has in fact the possibility of an energy transfer to Chl *a* that is as fast as the Chl *a* assembly equilibration, and the forms associated with these sites are exactly the closest in the Q_y region.

Among the many pigments present in the CP29 complex, Chl A2 and A5 form van der Waals contacts with Lutein 1 and 2, respectively, and are directly facing their central region (Figure 2). The 2-fold symmetry axis existing in LHCII, conserved in our model of CP29, makes the position of A2 on Lut 1 extremely similar to that of A5 on Lut 2 (the superimposition of the two pairs gives a root-mean-square deviation of 0.62 Å). Significantly, the Chl rings and carotenoid (Car) polyene chains are coplanar with an angle between their TMVs ranging from 12° (A2 to Lut 1) to 19° (A5 to Lut 2). All these features clearly suggest a role of these two groups in the carotenoid to chlorophyll energy transfer in light of the lowest-energy transition attributed, by LD reconstruction, to Chl A2 and A5 Q_y transitions. About the mechanism by which this transfer can take place, the simultaneous presence of a short distance and parallel TMV implies that both Dexter excitation exchange or excitation exchange through charge transfer states and strong Coulomb coupling may be active. For the first mechanism, only the pigment contact is necessary, and the independence of fluorescence yield allows the excitation exchange between the Car symmetry-forbidden first excited state and Chl Q_y transition. The experimental observation of this process has been reported in the literature by Trautman et al. (48), and more recently reviewed for bacterial photosynthesis by Koyama et al. (49). Strong Coulomb interaction should dominate anyway but through the Car second excited state, which is the only symmetry-allowed transition from the Car ground state. Whether this channel acts in CP29 as a part of light harvesting or another function cannot be determined here; however, the extremely short distance and the almost parallel orientation of TMVs should make the energy transfer in the two Chl–Car pairs extremely fast. We stress that the two Chls involved here, A2 and A5, belong separately to the two most intraconnected (and interconnected) Chl *a* pairs, A1–A2 and A4–A5; this guarantees the most effective equilibration of excitation within the whole pigment–protein complex.

REFERENCES

- Bassi, R., and Dainese, P. (1992) *Eur. J. Biochem.* 204, 317–326.
- Boekema, E., van Roon, H., Calkoen, F., Bassi, R., and Dekker, J. (1999) *Biochemistry* 38, 2233–2239.

3. Harrer, R., Bassi, R., Testi, M. G., and Shaefer, C. (1998) *Eur. J. Biochem.* 255, 196–205.
4. Jennings, R. C., Bassi, R., Garlaschi, F. M., Dainese, P., and Zucchelli, G. (1993) *Biochemistry* 32, 3203–3210.
5. Zucchelli, G., Dainese, P., Jennings, R. C., Breton, J., Garlaschi, F. M., and Bassi, R. (1994) *Biochemistry* 33, 8982–8990.
6. Kuhlbrandt, W., Wang, D. N., and Fujiyoshi, Y. (1994) *Nature* 367, 614–621.
7. Haworth, P., Arntzen, C. J., Tapie, P., and Breton, J. (1982) *Biochim. Biophys. Acta* 682, 152–159.
8. Hemelrijk, P. W., Kwa, S. L. S., van Grondelle, R., and Dekker, J. P. (1992) *Biochim. Biophys. Acta* 1098, 159–166.
9. Kuhlbrandt, W., and Wang, D. N. (1991) *Nature* 350, 130–135.
10. Dainese, P., and Bassi, R. (1991) *J. Biol. Chem.* 266, 8136.
11. Plumley, F. G., and Schindt, G. W. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 146–150.
12. Paulsen, H., Rümmler, U., and Rüdinger, W. (1990) *Planta* 181, 204–211.
13. Cammarata, K. V., and Schmidt, G. W. (1992) *Biochemistry* 31, 2779–2789.
14. Giuffra, E., Cugini, D., Croce, R., and Bassi, R. (1996) *Eur. J. Biochem.* 238, 112.
15. Ros, F., Bassi, R., and Paulsen, H. (1998) *Eur. J. Biochem.* 253, 653–658.
16. Pagano, A., Cinque, G., and Bassi, R. (1998) *J. Biol. Chem.* 273, 17154–17165.
17. Walters, R. G., Ruban, A. V., and Horton, P. (1994) *Eur. J. Biochem.* 226, 1063–1069.
18. Pesaresi, P., Sandonà, D., Giuffra, E., and Bassi, R. (1997) *FEBS Lett.* 402, 151–156.
19. Green, B. R., Pichersky, E., and Kloppstech, K. (1991) *Trends Biochem. Sci.* 16, 181–186.
20. Giuffra, E., Zucchelli, G., Sandonà, D., Croce, R., Cugini, D., Garlaschi, F. M., Bassi, R., and Jennings, R. C. (1997) *Biochemistry* 36, 12984–12993.
21. Bassi, R., Croce, R., Cugini, D., and Sandonà, D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*, Vols. 1–3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Gibson, T. J. (1984) Ph.D. Thesis, Cambridge University, Cambridge, England.
24. Gottesman, S., Halpern, E., and Trisler, P. (1981) *J. Bacteriol.* 148, 265–273.
25. Yukenberg, P. D., Withey, F., Geisselsoder, J., and McClary, J. (1991) in *Directed mutagenesis, A practical approach* (McPherson, M. J., Ed.) pp 27–48, IRL Press, Oxford, U.K.
26. Sanger, F., Nicken, S., and Carlson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5468.
27. Nagai, K., and Thogersen, H. C. (1987) *Methods Enzymol.* 153, 461–481.
28. Dainese, P., Hoyer-Hansen, G., and Bassi, R. (1990) *Photochem. Photobiol.* 51, 693–703.
29. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) *Anal. Biochem.* 150, 76–85.
30. Hirs, C. H. W. (1967) *Methods Enzymol.* 11, 325–329.
31. Porra, R. J., Thompson, W. A., and Kriedemann, P. E. (1989) *Biochim. Biophys. Acta* 975, 384–394.
32. Gilmore, A. M., and Yamamoto, H. Y. (1991) *J. Chromatogr.* 543, 137–145.
33. Breton, J., Michel-Villaz, M., and Paillot, G. (1973) *Biochim. Biophys. Acta* 314, 42–56.
34. Holm, L., and Sander, C. (1991) *Mol. Biol.* 218, 183–194.
35. Bassi, R., and Simpson, D. J. (1986) *Carlsberg Res. Commun.* 51, 363–370.
36. Koepke, J., Hu, X., Muenke, C., Schulten, K., and Michel, H. (1996) *Structure* 4, 581–597.
37. McDermott, G., Prince, S. M., Freer, A. A., Hawthorthwaite-Lawless, A. M., Papiz, M. Z., Cogdell, R. J., and Isaacs, N. W. (1995) *Nature* 374, 517–521.
38. Fragata, M., Norden, B., and Kurucsev, K. (1988) *Photochem. Photobiol.* 47, 133–143.
39. Van Gorp, M., van Langen, H., van Ginkel, G., and Levine, J. K. (1988) *J. Theor. Biol.* 131, 333–349.
40. Sandonà, D., Croce, R., Pagano, A., Crimi, M., and Bassi, R. (1998) *Biochim. Biophys. Acta* 1365, 207–214.
41. Zucchelli, G., Garlaschi, F. M., and Jennings, R. C. (1996) *Biochemistry* 35, 16247–16254.
42. Ide, J. P., Klug, D. R., Kuhlbrandt, W., Giorgi, L. B., and Porter, G. (1987) *Biochim. Biophys. Acta* 893, 349–364.
43. Bassi, R., Rigoni, F., and Giacometti, G. M. (1990) *Photochem. Photobiol.* 52, 1187–1206.
44. Peterman, E. J. G., Dekker, F. M., van Grondelle, R., and Van Amerongen, H. (1995) *Biophys. J.* 69, 2670–2678.
45. Gradinaru, C. C., Ozdemir, S., Gulen, D., van Stokkum, I. H. M., van Grondelle, R., and van Amerongen, H. (1998) *Biophys. J.* 75, 3064–3077.
46. Gradinaru, C. C., Pascal, A., van Mourik, F., Robert, B., Horton, P., van Grondelle, R., and van Amerongen, H. (1998) *Biochemistry* 37, 1143–1149.
47. Visser, H. M., Kleima, F. J., van Stokkum, R., van Grondelle, R., and van Amerongen, H. (1996) *Chem. Phys.* 210, 197–312.
48. Trautman, J. K., Shreve, A. P., Violette, C. A., Frank, H. A., Owens, T. G., and Albrecht, A. C. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 215–219.
49. Koyama, Y., Michitaka, K., Andersson, P. O., and Gillbro, T. (1996) *Photochem. Photobiol.* 63 (3), 243–256.