Localization of the N-terminal Domain in Light-harvesting Chlorophyll $a/b$ Protein by EPR Measurements*

Received for publication, February 1, 2005, and in revised form, March 8, 2005

Published, JBC Papers in Press, March 8, 2005, DOI 10.1074/jbc.M501171200

Gunnar Jeschke‡, Alexander Bender§, Thorsten Schweikardt¶, Grazyna Panek‡, Heinz Decker¶, and Harald Paulsen§

From the ‡Max-Planck-Institut für Polymerforschung, Postfach 3148, 55021 Mainz, §The Institut für Allgemeine Botanik der Johannes-Gutenberg-Universität Mainz, Müllerweg 6, 55099 Mainz, and the ¶Institut für Molekulare Biophysik der Johannes-Gutenberg-Universität Mainz, Wengerweg 6, 55128 Mainz, Germany

The conformational distribution of the N-terminal domain of the major light-harvesting chlorophyll $a/b$ protein (LHCIIb) has been characterized by electron-electron double resonance yielding distances between spin labels placed in various domains of the protein. Distance distributions involving residue 3 near the N terminus turned out to be bimodal, revealing that this domain, which is involved in regulatory functions such as balancing the energy flow through photosystems (PS) I and II, exists in at least two conformational states. Models of the conformational sub-ensembles were generated on the basis of experimental distance restraints from measurements on LHCIIb monomers and then checked for consistency with the experimental distance distribution between residues 3 in trimers. Only models where residue 3 is located above the core of the protein and extends into the aqueous phase on the stromal side fit the trimer data. In the other state, which consequently is populated only in monomers, the N-terminal domain extends sideways from the protein core. The two conformational states may correspond to two functional states of LHCIIb, namely trimeric LHCIIb associated with PSII in stacked thylakoid membranes and presumably monomeric LHCIIb associated with PSI in nonstacked thylakoids. The switch between these two is known to be triggered by phosphorylation of Thr-6. A similar phosphorylation-induced conformational change of the N-terminal domain has been observed by others in bovine annexin IV which, due to the conformational switch, also loses its membrane-aggregating property.

The major light-harvesting chlorophyll $a/b$ complex (LHCIIb)$^3$ in higher plants contributes to photosynthesis most significantly by increasing the amount of light energy absorbed and thus of the energy flow through the photosynthetic reaction centers. This is accomplished by the pigments present in the complex, eight chlorophyll (Chl) $a$ and six Chl $b$ molecules as well as four carotenoids per apoprotein, which are noncovalently bound to the protein and positioned such that the excitation energy can rapidly be conducted toward the reaction centers. With regard to this function, it seems reasonable to view the LHCIIb apoprotein as a rigid scaffold keeping the pigments in their proper position.

In addition to its light-harvesting functions, LHCIIb is involved in a number of regulatory processes that may require a more flexible behavior of the apoprotein. One of these processes ensures a balanced energy flow through the two photosystems (PS). When PSII reduces more plastoquinone than can be re-oxidized by PSI, then LHCIIb becomes phosphorylated at a threonine residue close to the N terminus by a thylakoid kinase. This modification prompts LHCIIb to dissociate from PSIII and bind to PSI, increasing the absorption cross-section of the latter (1, 2). The structural basis for this switch in binding partners is unclear, but an altered arrangement of the N-terminal hydrophilic domain in LHCIIb, triggered by its phosphorylation, has been discussed (3). Phosphorylation of LHCIIb has been reported to be favored not only by a reduced plastoquinone pool but also, more directly, when LHCIIb itself is excited by light. Irradiation with visible light causes an increased accessibility of the N-terminal protein domain to both the kinase and proteases (4). This apparent structural change, which is reversible in the dark, has been observed with thylakoids preparations and also with isolated LHCIIb, and therefore appears to be an intrinsic property of the pigment-protein complex (5). Moreover, LHCIIb is involved in the protection of the photosynthetic apparatus against photoinhibition due to overexcitation. Under high light conditions, one of the carotenoid components of the Chl $a/b$ antenna of PSI, violaxanthin, is converted to zeaxanthin by a de-epoxidase residing in the lumen, and this conversion, together with the acidification of the lumen compartment, enables the light-harvesting apparatus to dissipate excitation energy into heat rather than funneling it to the reaction centers. This quenching process requires the PSII protein PsbS (6) and is most likely brought about by a Chl-zeaxanthin heterodimer undergoing charge separation (7). The structural basis of this regulatory process is still unclear, but one of the explanations that have been suggested is an allosteric mechanism leading to a dissipating conformation of Chl $a/b$ complexes (8, 9) and possibly inducing a partial LHCIIb trimer-monomer transition (10).

In order to verify and understand the proposed structural changes in LHCIIb, the conformational behavior of the complex under various physiologically significant conditions needs to be analyzed. One possibility of doing so is to measure intramolecular distances. For instance, if structural transitions of the N-terminal domain are to be characterized, distances between the N terminus and other points in the molecule should be...
measured. Förster resonance energy transfer has been extensively used to measure molecular distances and specifically to characterize conformational changes in proteins (11, 12). The drawback of Förster resonance energy transfer for analyzing LHCIIb conformations is that, in order to avoid spectral interference with the intrinsic pigments, one would have to use fluorescent labels in the infrared spectral domain, which tend to be bulky and are likely to distort the protein structure to be studied.

An alternative technique for measuring molecular distances that avoids this complication is EPR. Spin labels used for EPR, such as tetramethylpiperidine-1-oxyl (TEMPO), compare in size to amino acid side chains and therefore are less likely to interfere with the structural behavior of the protein. Distances between such labels up to 20 Å can be measured with good precision by conventional continuous-wave EPR techniques (13). The combination of such measurements with site-directed spin labeling has recently emerged as a viable approach for studying structure and conformational dynamics of membrane proteins and protein complexes (14–16). The distance range of this approach can be extended up to at least 50 Å by applying pulse EPR techniques such as pulse electron-electron double resonance experiments (17, 18) or double-quantum EPR (19).

In particular, the four-pulse DEER experiment has been shown to be applicable to membrane proteins even if distances are rather broadly distributed as is the case for spin labels that are situated in loops (20). In recent methodological studies on extraction of distance distributions from four-pulse DEER data (21) and sensitivity enhancement (22), we have used doubly spin-labeled samples of LHCIIb as model systems. By applying double-quantum EPR to the soluble protein lysozyme T4 as a model system, Bobar et al. (23) demonstrated that EPR distance measurements combined with triangulation techniques can potentially be used to determine structural features of a protein.

In the present work, we attempted for the first time to obtain information on unknown structural details of a membrane protein by the combination of EPR distance measurements and triangulation. In particular, we try to determine possible locations of residue 3 close to the N terminus, which is not resolved in the crystal structures. On the basis of bimodal experimental distance distributions, we assume a two-state model of the N-terminal domain. Conformations of the domain that are consistent with the different possible assignments of the peaks in the distance distributions for LHCIIb monomers are modeled by using the program MODELLER, and it is tested as to which assignments are also consistent with the experimental data for the trimers. The results are discussed in the context of regulation of the energy flow balance of the plant photosystems.

MATERIALS AND METHODS

Protein Preparation and Spin Labeling—For the present study, two types of Lhcb1 clones derived from the Lhcb1*2 (AB80) gene (24) from pea (Pisum sativum) were used. Mutant versions of the protein were constructed containing a single Cys replacing Ser in position 3 (S3C) and Ser-160 in the stromal loop (S160C). Double mutants contained two cysteines in positions 3 and 160 (S3C,S160C), 3/52 (S3C,S52Ch), 52/160 (S52C,S160C), 106/160 (S106C,S160Ch), 3/106 (S3C,S160Ch), and 106/229 (S106C,S229C). In all these mutants, the native single Cys in positions 3 and 160 (S3C,S160C), 3/52 (S3C,S52Ch), 52/160 (S52C,S160C), 106/160 (S106C,S160Ch), 3/106 (S3C,S160Ch), and 106/229 (S106C,S229C). In all these mutants, the native single Cys in position 79 was replaced by Ser and, except for mutants 3/106 and 106/229, a hexahistidyl (His6) tag (designated by “H”) was added to the C terminus.

All mutants were overexpressed in Escherichia coli as described previously (25). The resulting proteins were purified (1 mg/ml) in an aqueous solution of 0.5% (w/v) SDS, 20 mM sodium phosphate (pH 7), and 2 mM tris-(2-carboxyethyl)phosphine and were incubated for 2 h at room temperature. Spin labeling was performed by adding 4-(2-iodoacetoamide)-2,2,6,6-tetramethylpiperide-1-oxyl (Sigma, 10-fold molar excess over protein) and incubating overnight at ambient temperature on a shaker. The proteins were then precipitated by adding 1 M acetic acid to a final concentration of 100 mM and 2.3 volumes of acetone. The protein was collected by centrifugation (12,000 × g for 60 min at 4 °C), washed several times with 70% ethanol and once with ethanol, and dried for 15 min at ambient temperature.

Preparation of Pigments and Recombinant LHCIIb—Chls and carotenoids were isolated from pea thylakoids according to Ref. 26. From all doubly labeled apoproteins, monomeric pigment-protein complexes were reconstituted as described previously (27). In addition, two monomeric apoproteins containing cysteine instead of serine at positions 3 (S3C) and 160 (S160C) were reconstituted. Trimerization of monomeric LHCIIb was initiated by immobilization of the His-tagged monomers on a nickel-chelating fast-flow Sepharose (Amersham Biosciences) in columns of 0.8 cm width and 4 cm length. The column was washed with 1 volume of OG buffer containing n-octyl glycoside (1% (w/v)), 0.1 M Tris-HCl (pH 9.0), 12.5% sucrose (w/v%). Monomeric LHCIIb was applied, incubated for 30 min at 4 °C, and washed with 3 column volumes of OG buffer, followed by 2 volumes of TX buffer, containing 0.05% Triton X-100 (v/v), 0.1 M Tris-HCl (pH 7.5), and 0.1 mg/ml dipalmitoyldiarylsulfonphoglycerol (Avanti Polar Lipids). Pigment-protein complexes were eluted from the column by addition of 1 volume of eluate buffer (0.05% Triton X-100 (v/v), 10 mM Tris-HCl (pH 7.5), 0.3 mM imidazole, and 0.1 mg/ml dipalmitoyldiarylsulfonphoglycerol). Monomeric and trimeric LHCIIb in the nickel-chelating column eluate were purified by sucrose density gradient ultracentrifugation. The eluate was loaded on to 1.01 x 1.02 density gradient containing 0.1% (w/v) lauryl maltoside, 5 mM Tricine (pH 7.8). After centrifugation for 17 h at 230,000 × g and at 4 °C, the bands containing monomeric and trimeric LHCIIb were collected and concentrated by Centricon centrifugal filter units (30 kDa, Bio-Rad) to 400 μM protein. Before loading to EPR tubes, the samples were mixed at a 1:1 volume ratio with 20% glycerol as a cryoprotectant.

EPR Measurements—Dipolar time evolution data were obtained at X-band frequencies (9.3–9.4 GHz) with a Bruker Elexys 580 spectrometer equipped with a Bruker Flexline split-ring resonator ER 4118X_MS3. Microwave from a YIG oscillator (Avantek AV 78012) customized by Magnettech GmbH (Berlin, Germany) was fed into one microwave pulse-forming unit of the spectrometer to provide the pump pulses. All measurements were performed using a DEER experiment: τ2(νpump) – τ1 – π(νpump) – τ’ – π(νpump) – (τ1 + τ2 – τ’) – π(νpump) – τ1 – echo (18). Time τ is varied, whereas τ1 and τ2 are kept constant, and the dipolar evolution time is given by t = τ’ – τ1. Data were analyzed only for t > 0. The resonator was overcoupled to Q = 100; the pump frequency νpump was set to the center of the resonator dip and coincided with the maximum of the nitroxide EPR spectrum, whereas the spin frequency νS was 65 MHz higher and coincided with the low field local maximum of the spectrum. All measurements were performed at a temperature of 50 K with observer pulse lengths of 32 ns for both π/2 and π pulses and a pump pulse length of 12 ns. Proton modulation was averaged by adding traces at eight different τ1 values, starting at τ1 = 200 ns and incrementing by Δτ = 8 ns. The total measurement time for each sequence was between 50 and 100 s.

Data Analysis—The theoretical model for data analysis is described in Refs. 21 and 28. Distance distributions were obtained by least squares fitting of amplitudes at a small number (9, 17, or 33) of equidistant sampling points in the range between 1.75 and 8 nm and interpolating between sampling points by cubic Hermite polynomials (21). A home-written Matlab program for this task is freely available on request. Similar distance distributions were obtained by Tikhonov regularization in the same distance range using a regularization parameter of 500.

Modeling of the N-Terminal Region of the LHCIIb—The program Modeller 6, version 2 (29), was used to generate structural models of LHCIIb on the basis of the recently solved crystal structure of the trimeric complex from spinach (30). MODELLER performs comparative protein structure modeling by satisfying spatial restraints. The program optimizes the structure of homologs by modeling a global probability density function that integrates stereochemical parameters and homology-derived restraints (31). The LHCIIb x-ray crystal structure employed as a template defines the positions of amino acids 14–231; therefore, the arrangement of the first 13 amino acids and the C terminus in the crystal model. As inputs, the intramolecular distances measured in this work between the spin labels in position 3 and in another position were used. For each of distances, 500 independent modeling procedures were performed to cover all roughly equally possible arrangements of the N-terminal protein domain that are consistent with the intramolecular distances measured. The variability in the models was ensured by starting out from slightly different structures. Modeller generated 500 identical
RESULTS

EPR Measurements of Doubly Spin-labeled LHClIb Monomers—Six different intramolecular distances within an LHClIb monomer were assessed by pulse EPR measurements. For each distance, a doubly spin-labeled Lhcb1 was prepared, carrying two TEMPO labels in amino acid positions where a serine (positions 3, 52, 106, and 160) or a valine (position 229) had been replaced with a cysteine (Fig. 1). The labeling positions (given in parentheses) were close to the N terminus (position 3), and in the N-terminal domain were close to the first transmembrane helix (position 52), in the stromal (position 106) and luminal loops (position 160), and close to the C terminus (position 229). The only cysteine in the native Lhcb1 amino acid sequence had been exchanged with a serine, so that labeling with an excess amount of the sulphydryl-reactive iodoacetamide derivative of TEMPO occurred exclusively at the intended positions. As a control, an Lhcb1 mutant containing no cysteines gave no EPR signal after it had been subjected to the same labeling procedure as the cysteine-containing mutants (data not shown). The efficiency of TEMPO labeling in Lhcb1 mutants containing a singular cysteine was quantified by subsequent fluorescence labeling with a 10-fold molar excess of BODIPY 507/515 IA (Molecular Probes D-6004). The fluorescence labeling dropped to less than 10% when the protein had been reacted previously with the TEMPO label, indicating that TEMPO labeling was at least 90% efficient (not shown). This means that at least 80% of the protein molecules used in the intra-molecular distance measurements carried two labels.

The spin-labeled proteins were reconstituted with pigments in detergent solution, and the resulting LHClIb derivatives were isolated by ultracentrifugation through a sucrose density gradient and concentrated by ultrafiltration to $>400$ μM. That this preparation contained monomeric LHClIb was verified by virtually complete energy transfer from Chl b to Chl a (25) and by circular dichroism spectra in the visible range (33) (data not shown).

Pulse EPR measurements with doubly spin-labeled LHClIb are exemplified in Fig. 2 for the mutant S3C/S160C, covering the distance between the N terminus and the stromal loop. The four-pulse DEER data (solid line in Fig. 2A) were corrected for the background from interactions with labels on other protein molecules (dashed line) and used to calculate distance distributions by cubic Hermite interpolation between sampling points (21). As a control, we performed a measurement on a 1:1 mixture of LHClIb monomers carrying only a single spin label in position 3 or 160, respectively, at approximately the same total protein concentration. The DEER decay in this case corresponds to the background from interactions with labels on other protein molecules (dots in Fig. 2A). The signals of bilabeled samples were corrected for this background contribution and, thus, correspond to distances between labels within the same LHClIb molecule.

Fig. 3 shows the distance distributions obtained with all six different spin label pairs. In general, primary DEER data from different preparations of the same double mutant (monomers) or single mutant (trimers) were tested for significant differences after normalization of the modulation depth as described previously (21). The primary data sets for different preparations of the same mutant (17 data sets for S3C/S160C, 12 data sets for S106C/S160C, and 2 data sets for each of the other mutants) were found to agree within experimental noise. They were at first processed independently, and for each of them the root mean square error for the best fit distance distribution was determined. To obtain an average primary data set, the data sets for a given mutant were weighted with their inverse root mean square error, added, and normalized. The distance distribution was then determined from the average primary data set. This procedure corresponds to a maximum likelihood estimate of the distance distribution from the series of independent primary data sets.

When we assume a fixed distance between the C atoms of the two residues, the maximum possible peak width in the distance distribution is determined by the conformational distribution of the spin labels and can be estimated as $\pm 0.5$ nm (13). Widths of the distribution larger than that imply a distri-
Localization of the N-terminal Domain in LHCIIb by EPR

FIG. 3. Distance distributions for six intramolecular distances in LHCIIb monomers. The location of the spin label pairs in the sequence is given in the upper left corners of the plots. Arrows indicate the distances taken from the crystal structure (30) between the midpoints of the amino acid side chains in the labeling positions.

bution of the distance of the Cα atoms. Shallow valleys in the distribution, as the one observed for mutant 106/160 at 3.7 nm, may be noise-induced and should be ignored. For mutant 3/160, where data could be measured only for maximum dipolar evolution times of 1.5 μs, we may not exclude the presence of distances larger than 4.5 nm, which would be suppressed by background correction for this evolution time. In the other cases, such contributions could be excluded by measuring data up to longer maximum dipolar evolution times and checking that the distance distribution did not change.

Three of the distances measured can be compared with the corresponding distances in the LHCIIb crystal structure (arrows in Fig. 3). The values taken from the crystal structure for distances 106/160 (3.89 nm) and 106/229 (3.50 nm) coincide quite well with the maxima of the distance distributions measured by EPR (3.86 and 3.84 nm, respectively). The distance 52/160 from the crystal structure (3.1 nm) is on the lower side of but still being comprised in the distance distribution obtained here (see also arrows in Fig. 3). Taken together, the data obtained by pulse EPR are fully consistent with the crystal structure. We can therefore conclude that the labeling of the protein with TEMPO does not disrupt the native protein structure. The EPR distance distributions including position 3 near the N terminus cannot be compared with the crystal structure in which the 13 N-terminal amino acids are not resolved.

Three of the distance distributions are bimodal, each showing two separated peaks centered at 2.8 and 4.1 nm (3/52), 2.9 and 4.0 nm (3/160), and 3.8 and 4.7 nm (52/160). In particular, the data for 3/52, where base-line separation of the peaks is at least as broad as the peaks themselves, strongly suggests that the N-terminal domain can exist in two different states. The remaining distributions are either broader monomodal or polymodal ones with less distinct peaks. The distribution of distances between the N terminus and the luminal loop (3/106) covers the range of 4–6.5 nm with a peak at 5.7 nm. This is at the upper limit of distances that can be reliably measured by pulse EPR and, therefore, prone to a larger error than the other distances measured. The separation between the two peaks in the 52/160 distribution is smaller than 1 nm and therefore could be due to different orientations of the spin labels. In the other two bimodal distance distributions, the peaks are further apart. Therefore, at least the data for distances 3/52 and 3/160 indicate a heterogeneous population of conformers in LHCIIb in solution, exhibiting different structures of the N-terminal hydrophilic domain.

Intermolecular Distances in LHCIIb Trimer—Pulse EPR can also be used to measure intermolecular distances between residues in LHCIIb monomers within a trimeric complex. Using doubly labeled LHCIIb in its trimeric form would yield too many different distances between spin labels to be resolved. However, when the monomers are labeled only in a single position and if it is assumed that the three monomers in a trimer are arranged symmetrically, then a single label-label distance is expected in the absence of conformational distribution.

We used LHCIIb trimers uniformly labeled in either position 3 near the N terminus or in position 160 in the stromal loop of each monomer. In Fig. 4, the distances between these labeling positions are indicated by bars in a structural model of trimeric LHCIIb. The distance between positions 160 is expected, on the basis of the x-ray structure (30), to be 6.1 nm and thus is clearly outside the range of distances that can be assessed with reasonable precision. The measurement does show, however, that the distance distribution does not extend to distances shorter than 5.5 nm (Fig. 4), which is in full agreement with the crystal structure. An additional narrow peak at the upper limit of the distance range (8 nm, not shown) is an artifact that results from an imperfect correction of the intermolecular background due to labels in neighboring trimers. This background correction becomes more difficult at longer intramolecular distances, as the difference between typical intramolecular and intermolecular distances then becomes smaller.

The distance distribution obtained for labels in position 3 close to the N terminus of LHCIIb covers the range of 1.5–5.5 nm with peaks at 2.5 and 4.2 nm. Again, this cannot be compared with the crystal structure data because the N-terminal section of the protein has not been resolved. The apparently bimodal distance distribution may indicate a heterogeneous structure of LHCIIb trimers (or of the monomers within a trimer) exhibiting different arrangements of the flexible N-terminal domain. To exclude the possibility that peaks at distances longer than 5.5 nm were suppressed by background correction, primary DEER data were acquired up to an evolution time of 2.5 μs (Fig. 4B). As the decay of the signal for evolution times between 1 and 2.5 nm strongly resembles the decay in singly labeled monomers of this mutant, we can safely exclude that distances longer than 5.5 nm occur in trimers to a significant extent. Possible correlations between the trimer data and the distances obtained from LHCIIb monomers will be discussed below.

Modeling of the N-terminal Region of the LHCIIb Monomer—The distance distributions between position 3 and other amino acid positions in the LHCIIb molecule were used to obtain information on the relative position of the N terminus, which is not resolved in the crystal structures (30, 34). The software Modeler 6, version 2 (29), was employed to generate LHCIIb structures on the basis of the crystal structure of spinach LHCIIb (30) as a template. Consequently, only the N-terminal 13 amino acids could be freely positioned.
with the restraint that overlap and torsion conflicts were avoided. MODELLER was prompted to generate 500 different structures with physically possible arrangements of the N-terminal domain while incorporating the EPR-derived distances as additional restraints. Only distances 3/52 and 3/160 were directly used as restraints; distance 3/106 was initially disregarded because of the larger experimental error. Later checks showed, however, that all modeled structures had distances 3/106 that were within the experimental distribution (data not shown). Each structural model produced by the program resembled the template within 1.2 Å maximum deviation in all molecule sections that were resolved in the crystal structure. Because all the intramolecular distance distributions measured in this work between positions 52, 106, 160, and 229 included the corresponding distances seen in the crystal structure, all models are also consistent with these experimental EPR distance distributions.

The bimodal character of both the 3/52 and 3/160 distance distributions suggests that in LHCIIb in a fluid environment each of these distances preferentially adopts either a smaller or a larger value. For each of the possible combinations (short/short, short/long, long/short, long/long), 500 structural models were simulated restraining those distances to slightly larger intervals than were found experimentally. From these models we selected those ones that had distances within the experimentally found intervals (Table I).

From all 1514 monomer models, trimer models were constructed by fitting (DeepView) the simulated structure to each of the three molecules in the crystal structure (30) of the trimer. The distance distributions between the O' atoms of Ser-3 in the four ensemble models of the trimer are displayed in Fig. 5, in each case together with the experimental distance distribution (see also Fig. 4). Clearly, the two distance intervals for 3/52 (Fig. 5, A and B versus C and D) correspond to significantly different positions of residue 3, with only the shorter 3/52 distance being compatible with the 3/3 distance measured in the trimer. The situation is less clear cut for distance 3/160. Allowing for a shift in the distance distribution due to the measuring the distance between N-O groups in the labels rather than between the S' atoms of the cysteine residues, the modeled distribution in Fig. 5A (shorter distances for both 3/52 and 3/160) is a strikingly good fit of the experimental distribution. On the other hand, Fig. 5B shows that the combination of the shorter distance 3/52 with the longer distance 3/180 corresponds to a distribution of distance 3/3 that is also largely consistent with the measured data; therefore, the following discussion will include structural models in which both intervals are populated for 3/160. The ensemble model for the trimer structure thus corresponds to models fulfilling condition 1 or 2 in Table I (Fig. 5, A and B), whereas the ensemble model for a structure that exists only in monomers corresponds to conditions 3 or 4 (Fig. 5, C and D). Note that the latter type of "monomer-only" structure must exist, as otherwise we could not explain the prominent long distance peak in the distance distribution 3/52 in the absence of distances 3/3 longer than 5.5 nm in the trimer.

For visualization of the two ensemble models, we averaged structures of the residues found in the template (14–231) over the whole set of models. The N-terminal domains are represented only by the side group atoms of Ser-3 (shown as van der Waals surfaces), as we do not have experimental information on the structure of the loop between residues 3 and 14. For clarity, we selected only trimer models with distances 3/3 shorter than 5.2 nm and only monomer models with distances 3/3 longer than 6 nm (in each case 80% of all models). The ensemble models are presented in Fig. 6, A–D, as side views parallel to the membrane plane and as top views perpendicular to this plane from the stromal side.

In the ensemble structure of the trimer (Fig. 6, A and C), the N terminus is located above the core of the protein and extends into the aqueous phase. In contrast, in the monomer-only ensemble structure, the N terminus is located sideways from the protein core and may have contact with the lipid head groups. Note that the spatial distribution of residue 3 shown in Fig. 6, A–D, is likely to overestimate disorder of the N-terminal domain. This is because we did not distinguish between the two peaks in the distance distribution 3/160, which may well be correlated to the two states, and because any structural disorder of the stromal loop introduces additional uncertainty for distance 3/160 that blurs our estimate for the position of residue 3. Furthermore, we may expect that additional experimental distance restraints involving residue 3 would further narrow the ensemble of possible models. The very broad spatial distributions visualized in Fig. 6, A–D, are thus conservative estimates for all positions of residue 3 and are consistent with our data. Given the narrow peaks in the distance distribution 3/52, the actual distribution is most likely more confined.

**DISCUSSION**

**Interpretation of Distance Distributions and Structural Modeling**—Pulse-EPR has been used as a "spectroscopic ruler" to measure intra- and intermolecular distances in monomeric and trimeric LHCIIb, respectively. The measurements yield distance distributions exhibiting one or several peaks with peak widths at half-height between 0.5 and 2 nm. The actual end points of distance measurements were the nitroxide groups of the spin labels attached to the proteins. These nitroxide groups are estimated to be 0.4 nm away from the cysteine thiol to which the spin label is attached. Consequently, if the spin labels are in a fixed position with regard to the protein structure, the cysteine-cysteine separation may deviate by up to 0.8

---

| Condition no. | Distance 3/52 | Distance 3/160 | No. selected models |
|---------------|--------------|---------------|---------------------|
| 1             | 24–30        | 22–32         | 392                 |
| 2             | 24–30        | 38–45         | 346                 |
| 3             | 39–43        | 22–32         | 397                 |
| 4             | 39–43        | 38–45         | 379                 |
Localization of the N-terminal Domain in LHCIIb by EPR

Fig. 6. Two-state behavior of the N-terminal domains in LHCIIb (A–D) and bovine annexin IV (E and F). A, possible locations of residue 3 (gray balls, superposition of 729 models) in LHCIIb trimers (conditions 1 and 2 in Table I), side view parallel to the membrane plane. B, possible locations of residue 3 (gray balls, superposition of 616 models) that are only populated in LHCIIb monomers (conditions 3 and 4 in Table I), side view. C, possible locations of residue 3 in LHCIIb trimers, top view perpendicular to the membrane plane from the stromal side. D, possible locations of residue 3 that are only populated in LHCIIb monomers, top view. E, conformation of the N-terminal domain in wild-type bovine annexin IV (PDB code 1ANN), top view. F, conformation of the N-terminal domain in the "phosphorylation" mutant T6D of bovine annexin IV (PDB code 1I4A), top view. Residues 1–9 are not observed in this structure.

nm from the distances measured. In the more likely case that the spin labels possess some rotational freedom at their attachment site, this will add a signal width of up to 0.8 nm to the distance distributions. For most pairs of labels, the total width of the distribution is clearly larger than this variation as expected in a worst case scenario. The shape of the experimental distance distributions is therefore thought to reflect both this experimental uncertainty inherent in the labeling approach and the variability of the individual distances due to locally mobile protein domains or to the presence of different protein conformers in the sample.

Several of the distance distributions obtained in the pulse EPR measurements could be verified by comparison with the crystal structure of LHCIIb. An exception would be the measurements including position 3 near the N terminus because the N-terminal protein domain has not been resolved in the structural models derived from two- or three-dimensional crystal analysis (30, 34). In all cases that could be checked, the EPR distance data were consistent with the crystal structures in that the distances taken from the structures were within the distance distributions obtained by EPR. In the case of positions 106/160 (stromal loop/luminal loop) as well as 106/229 (luminal loop/C-terminal domain), the distances read from the structure even coincide with the peak values of the distributions (Fig. 3).

This is also true for the measurements with trimeric LHCIIb carrying spin labels in the luminal loop (position 160) of each monomer; however, this may be less significant than the two intra-molecular distances because this long distance of 6 nm is outside the range that can reliably be assessed by our present EPR setup.

On the other hand, distance 52/160 (N-proximal side of the first trans-membrane helix/stromal loop) in the crystal structure is almost 1.5 nm shorter than the peak in the EPR-measured distance distribution. This large difference and the extension of the distance distribution to even larger distances clearly indicate a structural difference between solubilized LHCIIb and the crystal structure. Because position 52 is part of the presumably more rigid α-helical scaffold of the protein, the difference is likely to be confined to the stromal loop domain. This would also be consistent with the rather narrow peaks in the distance distribution 3/52 and the broader peaks in the distance distribution 3/160. However, our data are not consistent with the stromal loop extending far away from the stromal surface of the membrane, because this would result in a separation of the luminal and stromal loops (positions 106/160) significantly larger than that in the LHCIIb crystal, which we do not observe. Consequently, our data suggest that the stromal loop around position 160 in solubilized LHCIIb is still located close to the stromal surface of the molecule but on average shifted, in comparison to the crystal structure, further away from the N-terminal domain.

The distance distribution 52/160 exhibits two peaks with nearly base-line separation that are less than 1 nm apart. As conversion of the dipolar evolution function to a distance distribution is an ill-posed mathematical problem, narrow noise artifacts down to the base-line level can in principle occur (21). It is thus not immediately clear whether the distance distribution 52/160 is also bimodal. We therefore checked four measurements on two independently prepared samples using both Hermite interpolation between sampling points and Tikhonov regularization to compute the distance distributions. The minimum close to 4 nm is observed in all data sets with both analysis methods. Hence, the distance distribution is very likely to be bimodal indeed. Together with the distance of 3.01 nm between residues 52 and 160 in the crystal structure, this suggests that the stromal loop may also exist in two conformations and that the one realized in the crystal structure is the less likely one in a fluid environment. Given that in the trimer the N-terminal loop is located above the protein core on the stromal side, it is unlikely that conformational switching of this domain influences the structure of the stromal loop. It is thus possible that only the conformation realized in the crystal structure exists in the trimer. However, our current experimental data are not sufficient to conclude that.

The two distance distributions 3/52 and 3/160 clearly are bimodal. Both contain two prominent peaks that are nearly base-line-separated and more than 1 nm apart. Note that in the latter case the distance distribution is an average from 17 independently prepared samples, each of them exhibiting this splitting in the distance distribution. We can thus safely conclude that the solubilized monomeric LHCIIb contains two approximately equally prominent populations of protein conformers that differ in their separations between the N terminus (position 3) and other positions on the stromal surface of the complex. The distances between the N terminus and a position on the luminal surface (3/106) are rather long (peak at >5 nm), so we cannot distinguish between whether this in fact is a single if broad distribution peak or actually a bimodal distribution that we cannot resolve due to experimental limitations. Modeling suggests that bimodality in this distribution would indeed be difficult to resolve, even if we could improve the distance range of our technique (data not shown).

As mentioned above, position 52 in the stromal extension of the N-proximal trans-membrane α helix is likely to be part of the presumably more rigid α-helical protein frame. The bimodal distance distribution 3/52 therefore means that the hydrophilic N-terminal domain in solubilized LHCIIb adopts at least two significantly different conformations. This conformational flexibility may be of functional significance because the N-terminal domain is involved in the structural and regulatory behavior of the complex. A trimerization motif in positions 16–21 is essential for LHCIIb trimer formation (35); isolated trimeric LHCIIb has been reported to dissociate into monomers.
upon illumination (10) which, if it in fact occurs in vivo, may increase dissipation of excitation energy. Moreover, threonine phosphorylation in positions 5 and/or 6 triggers state 1–2 transition and the migration of LHCIib from PSII to PSI (1, 2). Light-induced structural changes in the N-terminal protein domain, as detected by altered accessibilities to protein kinase or protease, have been seen in LHCIib from pea (4) and Chlamydomonas reinhardtii (36), and conformational changes upon phosphorylation of Thr-5 have been observed in a synthetic peptide corresponding to the N terminus of LHCIib (3). The data presented here provide evidence, to our knowledge for the first time, that solubilized LHCIib in fact contains populations of at least two conformers exhibiting significantly different structures of the N-terminal domains.

LHCCIib in vivo is predominantly organized in trimeric complexes that are intrinsic to the thylakoid membrane. In order to verify that the LHCIib conformers observed are functionally significant, their presence would have to be shown in membrane-inserted LHCIib. These experiments are to be performed in our future work. It should be noted, however, that detergent micelles that are able to stabilize membrane protein complexes are thought to mimic their membrane environment quite well (37), and our knowledge about the LHCIib crystal structure also is based on complexes in which the lipid environment has largely been replaced by detergents.

The N terminus of LHCIib has not been resolved in either of the published crystal structures. This may also indicate that in the crystals the N-terminal domain adopted variable positions. Alternatively, the sequence heterogeneity between LHCIib approteins, which is mostly confined to the N-terminal protein sections (38), may be responsible for the lack of structural resolution. It is impossible to deduce the position of the N terminus when only single distances to other points in the molecule are known. However, on the basis of several such distances, it should in principle be possible to triangulate the unknown position in space. We made an attempt to localize the position of the N-terminal amino acids in LHCIib by calculating molecular models on the basis of the crystal structure of spinach LHCIib and using the EPR-derived distance ranges as restraints. The modeling procedure is complicated by the fact that the distance distributions between the N terminus and both the stromal extension of the first trans-membrane helix and the stromal loop (3/52 and 3/160, respectively) turned out to be bimodal. Because we do not know how the two peaks in each distance distribution correlate with each other, we modeled all four possible combinations (Table I). The third measured distance involving the N terminus and the luminal loop (3/106) was initially disregarded for modeling because this measurement in the range of ~5 nm was considered not to be precise enough. We did check, however, that all 1514 final models are consistent with our measurement of that distance, i.e. the O\(^\cdot\)-O\(^\cdot\) distance between residues 3 and 106 in all these structures is longer than 4.5 nm.

Distance distributions 3/3 in trimers constructed from the monomer models are consistent with the corresponding experimental distribution only for model families 1 and 2 that correspond to the peak at shorter distances in the distribution for 3/52. In the ensemble model of this trimer structure, the N terminus is located above the protein core (Fig. 6C) and extends into the aqueous phase on the stromal side (Fig. 6A). The peak at longer distances in the distribution for 3/52 must then correspond to conformations of the N-terminal domain that are populated in the monomer only. In this ensemble of structures, the N terminus is located sideways from the helix bundle (Fig. 6D) and possibly attaches to the membrane surface (Fig. 6B).

These conclusions depend on only one remaining assumption, namely that all conformations that are populated in the trimer are also accessible in the solubilized monomer. To prove this plausible assumption, measurements on a trimer composed of one doubly labeled LHCIib molecule and two unlabeled LHCIib molecules would have to be performed. Recent experiments on separation of LHCIib heterotrimers suggest that such experiments will be possible in the near future. If our current models are correct, we should then find a monomodal distance distribution 3/52 and possibly also monomodal distance distributions 3/160 and 52/160. If so, our models could be considerably refined.

The two states of the N-terminal domain found in the present work are very likely to be of functional significance. If dissociation of LHCIib trimers in vivo is a consequence of phosphorylation of Thr-5 as proposed by Allen and Forsberg (39), our model suggests that such phosphorylation causes the conformational change that moves the N-terminal domain from a position above the protein core to a position sideways from it. Rather intriguingly, a combined NMR and Fourier transformed infrared spectroscopy study of a peptide consisting of the 15 N-terminal residues of LHCIib and its counterpart phosphorylated at Thr-5 indicated a conformational switching of the N-terminal domain that is consistent with our picture (9). For the unphosphorylated peptide, Nilsson et al. (3) found a doubling of the signal of NH protons of Gly-13, with different contacts of the two forms to Ser-14. This suggests a bimodal conformational distribution, in agreement with the bimodal conformational distribution that we find for the N-terminal domain. The NMR results also showed that phosphorylation of Thr-5, which is featured in the regulation of LHCIib, influences this conformational switch at Gly-13 of the peptide.

**Analogy with Regulation of Annexins—**Much to our surprise, a rather similar conformational switching of an N-terminal domain has been implied in the regulation of annexin IV by phosphorylation of its residue Thr-6 (40). In the crystal structure of wild-type bovine annexin IV (PDB identifier 1ANN), the N-terminal domain is located above the core of the protein (Fig. 6E). The closest contact with the core is at Thr-6, which is

| Annexin | Organism         | PDB identifier | N-terminal site | Core site | Distance |
|---------|------------------|----------------|-----------------|-----------|----------|
| A2      | Homo sapiens     | 1W7B           | Ser-26          | Leu-334   | 2.52     |
| III     | H. sapiens       | 1AXN           | Thr-12          | Ile-319   | 2.58     |
| IV      | Bos taurus       | 1ANN           | Thr-6           | Leu-313   | 2.69     |
| V       | Gallus gallus    | 1ALA           | Thr-9           | Leu-315   | 2.57     |
| VI      | B. taurus        | 1AVC           | Ser-13          | Leu-320   | 2.56     |
| XII     | Hydra vulgaris   | 1AEI           | Thr-6           | Ile-313   | 2.94*    |
| Annexin II | Arabidopsis thaliana | 1WCN | Thr-3 | Leu-311 | 2.66 |
| Annexin IH | Gossypium hirsutum | 1N00 | Thr-8 | Leu-315 | 2.41 |
| Annexin 24 | Capsicum annuum | 1DK5 | Ser-11 | Leu-316 | 2.67 |

* Chain A of the hexamer.
hydrogen-bonded to the carbonyl oxygen of Leu-313. Inspection of crystal structures of annexins in the Protein Data Bank reveals that such a location of the N-terminal domain above the core with a threonine or serine residue at hydrogen-bonding distance to a carbonyl oxygen in the core is conserved between different annexins with short and intermediate N termini from both animals and plants (Table I). This is despite otherwise substantial variations in sequence and even in length of the N-terminal domain. Kaetzel et al. (40) mimicked phosphorylation of Thr-6 in bovine annexin IV by mutation Thr to Asp. The crystal structure of T6D annexin IV (PDB identifier 1I4A) confirms that with the introduction of aspartate at this position the N-terminal domain is released from the rft on the surface of the protein core (Fig. 6F), just as we found for the monomer-only conformation of the N-terminal domain of LHCIIB. In bovine annexin IV, Gly-12 acts as a hinge, whereas in LHCIIB both Gly-13 and Pro-15 may be involved. In fact, many of the annexins listed above have glycine or proline residues close to the putative hinge.

These findings point to an analogous hinge motion of the N-terminal domain in annexins and LHCIIB, which in both cases is probably preceded by phosphorylation of a threonine or serine residue as the previous step of a regulation cascade. In fact, the next step might also be analogous, as annexins are known to aggregate membranes (41), and the trimer state of LHCIIB has been implied in thylakoid membrane stacking in the grana (39). It is known that phosphorylation of Thr-6 in bovine annexin IV impairs its function in membrane aggregation (40). This feature is shared by phosphorylation of the analogous residues in other members of the annexin family of proteins (41). In the case of annexin A1, there is some evidence that the unphosphorylated N-terminal domain provides a secondary site for membrane attachment. It is also known that only the LHCIIB trimer associates with photosystem II in grana, whereas the monomer resides in stroma that are unstacked thylakoids (39).

In conclusion, we have demonstrated that it is feasible to triangulate the position of a residue in a flexible domain of a membrane protein by EPR measurements. For proteins whose regulation depends on switching between two clearly distinct conformational states, such triangulation can help to derive an hypothesis on the molecular basis of the switching. In particular, we have shown that solubilized LHCIIB contains populations of at least two conformers exhibiting significantly different structures of the N-terminal domains and that only one of these conformations is significantly populated in the trimer. The structural models for the two conformations exhibit an analogy to conformations observed in crystal structures of wild-type annexin IV and its mutant T6D, which may indicate an analogy in the regulation of these two proteins by phosphorylation. The conformational change in the N-terminal domain of LHCIIB may then cause not only a change in oligomerization state but also a change in the preference for stacked membrane regions.

Acknowledgment—We thank Martin Seimetz for help with some sample preparations.

REFERENCES
1. Allen, J. F. (2003) Science 299, 1530–1532
2. Zer, H., and Ohad, I. (2003) Trends Biochem. Sci. 28, 467–470
3. Nilsson, A., Styx, D., Drakenberg, T., Spangfort, M. D., Forsen, S., and Allen, J. F. (1997) J. Biol. Chem. 272, 18350–18357
4. Zer, H., Vink, M., Shochat, S., Herrmann, R. G., Andersson, B., and Ohad, I. (2008) Biochemistry 47, 728–738
5. Zer, H., Vink, M., Keren, N., Dilly Hartwig, H. G., Paulsen, H., Herrmann, R. G., Andersson, B., and Ohad, I. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 8277–8282
6. Li, X.-P., Björkman, O., Shi, C., Grossman, A. R., Rosenquist, M., Jansson, S., and Nyogi, K. (2000) Nature 403, 391–395
7. Holt, N. E., Ziemianinska, D., Valkunas, L., Li, X. P., Nyogi, K. K., and Fleming, G. R. (2005) Science 307, 433–436
8. Horton, P., Ruban, A. V., and Wentworth, M. (2000) Philos. Trans. R. Soc. Lond. B Biol. Sci. 355, 1361–1370
9. Wentworth, M., Ruban, A. V., and Horton, P. (2004) Biochemistry 43, 501–509
10. Garah, G., Cseh, Z., Kovacs, L., Rajagopalan, S., Varkonyi, Z., Holzenburg, A., and Horton, P. (2002) Biochemistry 41, 15121–15129
11. Tucker, M. J., Oyola, R., and Gai, F. (2005) J. Phys. Chem. B 109, 4788–4795
12. Johnson, A. E. (2005) FEBS Lett. 579, 916–920
13. Rabenstein, M. D., and Shin, K. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 8239–8243
14. Hubbell, W. L., and Altenbach, C. (1994) Curr. Opin. Struct. Biol. 4, 566–573
15. Hubbell, W. L., Cufeso, D. S., and Altenbach, C. (2000) Nat. Struct. Biol. 7, 735–739
16. Steinhoff, H. J. (2002) Front. Biosci. 7, 97–110
17. Milor, A. D., Salikhov, K. M., and Shirov, M. D. (1981) Fiz. Toreb. Telas 23, 973–982
18. Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W. (2000) J. Magn. Reson. 142, 331–340
19. Borbat, P. F., and Freed, J. H. (1999) Chem. Phys. Lett. 313, 145–154
20. Jeschke, G., Wegener, C., Nieschtag, M., Jung, H., and Steinhoff, H.-J. (1999) Biophys. J. 86, 2551–2557
21. Jeschke, G., Panek, G., Godt, A., Benda, A., and Paulsen, H. (2004) Appl. Magn. Reson. 26, 223–244
22. Jeschke, G., Bende, A., Paulsen, H., Zimmermann, H., and Godt, A. (2004) J. Magn. Reson. 169, 1–12
23. Borbat, P. F., Mchaourab, H. S., and Freed, J. H. (2002) J. Am. Chem. Soc. 124, 5309–5314
24. Cashmore, A. R. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 2960–2964
25. Paulsen, H., Rümker, U., and Rüdiger, W. (1990) Planta 181, 204–211
26. Booth, F. J., and Paulsen, H. (1998) Biochemistry 37, 5103–5108
27. Paulsen, H., Finkenzeller, B., and Kühlein, N. (1995) Eur. J. Biochem. 215, 809–816
28. Jeschke, G., Koch, A., Jonas, U., and Godt, A. (2002) J. Magn. Reson. 155, 72–82
29. Sali, A., and Blundell, T. L. (1993) J. Mol. Biol. 234, 779–815
30. Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gu, L., An X., and Chang, W. (2004) Nature 428, 287–292
31. Fiser, A., Do, R. K., and Sali, A. (2000) Protein Sci. 9, 1753–1773
32. Guex, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714–2723
33. Hobe, S., Prysulla, S., Kührbrandt, W., and Paulsen, H. (1994) EMBO J. 13, 3423–3429
34. Kührbrandt, W., Wang, D. N., and Fujisawa, Y. (1994) Nature 367, 614–621
35. Hobe, S., Förster, R., Königler, J., and Paulsen, H. (1995) Biochemistry 34, 10224–10228
36. Zer, H., Vink, M., Alumot, N., Gaathon, A., Niyogi, K., Herrmann, R. G., Anderson, B., and Ohad, I. (2004) Biochemistry 43, 7824–7833
37. Roth, M., Lewit-Bentley, A., Michel, H., Deisenhofer, J., Huber, R., and Oesterhelt, D. (1989) Nature 340, 659–662
38. Jansson, S., Pichersky, E., Bassi, R., Green, B. R., Ikeuchi, M., Melis, A., Simpson, D. J., Spangfort, M., Stachelin, L. A., and Thornber, J. P. (1992) Plant Mol. Biol. 19, 242–253
39. Allen, J. F., and Forsberg, J. (2001) Trends Plant Sci. 6, 317–326
40. Kaetzel, M. A., Mo, Y. D., Mealy, T. R., Campos, B., Bergsma-Schutter, W., Brissos, A., Dedman, J. R., and Seaton, B. A. (2001) Biochemistry 40, 4192–4199
41. Gerke, V., and Moss, S. E. (2002) Physiol. Rev. 82, 331–371