The Solution Structure of Photosystem I Accessory Protein E from the Cyanobacterium *Nostoc* sp. Strain PCC 8009[†,‡]

Kristen L. Mayer,[§,||] Gaozhong Shen,[⊥] Donald A. Bryant,[⊥] Juliette T. J. Lecomte,[*,§] and Christopher J. Falzone[§]

Department of Chemistry and Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802

Received May 6, 1999; Revised Manuscript Received August 13, 1999

ABSTRACT: PsaE is a small basic subunit located on the stromal (cytoplasmic) side of photosystem I. In cyanobacteria, this subunit participates in cyclic electron transport and modulates the interactions of the complex with soluble ferredoxin. The PsaE protein isolated from the cyanobacterium *Synechococcus* sp. strain PCC 7002 adopts the β topology of an SH3 domain, with five β strands (βA through βE) and a turn of 3_{10} helix between strands βD and βE [Falzone, C. J., Kao, Y.-H., Zhao, J., Bryant, D. A., and Lecomte, J. T. J. (1994) *Biochemistry 33,* 6052−6062]. The primary structure of the PsaE protein is strongly conserved across all oxygen-evolving photosynthetic organisms. However, variability in loop lengths, as well as N- or C-terminal extensions, suggests that the structure of a second representative PsaE subunit would be useful to characterize the interactions among photosystem I polypeptides. In this work, the solution structure of PsaE from the cyanobacterium *Nostoc* sp. strain PCC 8009 was determined by NMR methods. Compared to PsaE from *Synechococcus* sp. strain PCC 7002, this PsaE has a seven-residue deletion in the loop connecting strands βC and βD, and an eight-residue C-terminal extension. Angular and distance restraints derived from homonuclear and heteronuclear NMR experiments were used to calculate structures by a distance-geometry/simulated-annealing protocol. A family of 20 structures (rmsd of 0.24 Å in the regular secondary structure) is presented. Differences between the two cyanobacterial proteins are mostly confined to the CD loop region; the C-terminal extension is disordered. The thermodynamic stability of *Nostoc* sp. strain PCC 8009 PsaE toward urea denaturation was measured by circular dichroism and fluorescence spectroscopy, and thermal denaturation was monitored by UV absorption spectroscopy. Chemical and thermal denaturation curves are modeled satisfactorily with two-state processes. The $\Delta G°$ of unfolding at room temperature is 12.4 ± 0.3 kJ mol^{-1} (pH 5), and the thermal transition midpoint is 59 ± 1 °C (pH 7). Interactions with other proteins in the photosystem I complex may aid in maintaining PsaE in its native state under physiological conditions.

PsaE[1] is 1 of 11 polypeptides that comprise a monomeric unit of cyanobacterial photosystem I (PS I). The core of this complex is a heterodimer of the PsaA and PsaB proteins, containing the P700 reaction center, a 4Fe-4S cluster (denoted F_X), and most of the 100 chlorophyll *a*, 10−20 carotenoid, and 2 phylloquinone molecules of the monomeric unit (*1*). The functional complex is found along with photosystem II in all oxygen-evolving photosynthetic organisms, including higher plants, algae, and cyanobacteria. Cyanobacterial PS I complexes form trimers with a mass of approximately 1.05 MDa (*2*−*4*). A homodimeric reaction center with structural similarity to PS I is found in some anoxygenic bacteria, including the green sulfur bacteria and heliobacteria (*5*).

PS I catalyzes the light-driven transfer of an electron from reduced plastocyanin on the lumenal side of the thylakoid membrane to soluble, oxidized ferredoxin on the stromal side of the membrane. Reduced ferredoxin in turn produces NADPH through the action of the flavoprotein enzyme ferredoxin:NADP$^+$ oxidoreductase (FNR). Upon light absorption by PS I and energy migration to the special pair chlorophylls (P700), electron transfer occurs from the primary electron donor, P700, to the primary electron acceptor, A_0, which is a chlorophyll *a* molecule. The electron-transfer pathway then involves successive transfers to the intermediate acceptors A_1 (phylloquinone), the F_X cluster, and two 4Fe-4S clusters (F_A and F_B) bound by the protein

[†] This work was supported in part by National Institutes of Health Grant GM-54217 (to J.T.J.L.) and National Science Foundation Grant MCB-9723469 (to D.A.B.). G.S. was supported in part by National Science Foundation Grant DBI-9602232.

[‡] The coordinates have been deposited at the Protein Data Bank under the file names 1qp2 and 1qp3.

* To whom correspondence should be addressed at the Department of Chemistry, The Pennsylvania State University, 152 Davey Laboratory, University Park, PA 16802. E-mail: jtl1@psu.edu.

[§] Department of Chemistry.

[||] Present address: Department of Chemistry, Indiana University, Bloomington, IN 47405.

[⊥] Department of Biochemistry and Molecular Biology.

[1] Abbreviations: COSY, correlated spectroscopy; CT, constant time; 2D, two-dimensional; 3D, three-dimensional; FID, free induction decay; HSQC, heteronuclear single-quantum coherence; MD, molecular dynamics; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; PAGE, polyacrylamide gel electrophoresis; PDB, protein data bank; PS I, photosystem I; Psa, photosystem I subunit protein; 2Q, two-quantum; 2QF, double-quantum filter; SDS, sodium dodecyl sulfate; rmsd, root-mean-square deviation; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase incrementation; Tris, tris(hydroxymethyl)aminomethane.

10.1021/bi9910373 CCC: $18.00
Published on Web 09/23/1999

PsaC, before ending with the 2Fe-2S cluster of ferredoxin. Electrons can also travel from ferredoxin back to $P700^+$ in a process known as cyclic electron transport. Cyclic electron transport additionally requires the participation of the cytochrome b_6f complex, plastocyanin, or cytochrome c_6 and possibly other components [e.g., a type 1 NADH dehydrogenase (*6*)]. When this path is taken, NADPH cannot be produced, but the electron transport can be utilized to generate a protonmotive force usable for ATP synthesis.

PsaC, PsaD, and PsaE are extrinsic proteins found on the stromal (cytoplasmic) surface of PS I (*7*, *8*). PsaD stabilizes the binding of PsaC onto the complex and provides a docking site for ferredoxin (*9*). The role of PsaE has been elucidated with PS I complexes from which the subunit has been removed (*10*−*13*). These PsaE-less complexes revealed that the PsaE subunit is not essential to the organism's survival under normal growth conditions (*14*, *15*). However, PsaE-less cells grow slowly under stress conditions such as low light intensity or low carbon dioxide concentrations, or both (*13*). It was found that in those cells, the kinetics of $P700^+$ reduction are significantly slower, an indication that the PsaE subunit is required for cyclic electron transport around PS I (*16*).

PsaE is also involved in the interaction of PS I with ferredoxin or flavodoxin (*17*, *18*). Flavodoxin is a ferredoxin substitute that is typically synthesized by cyanobacteria under conditions of iron limitation (*19*, *20*). Cross-linking between the PS I complex and either one of these two electron acceptors is possible in the wild-type complex, but cross-linking is reduced and altered in PS I complexes that do not contain PsaE (*18*). In addition, once cross-linking has been established in the wild-type complex, PsaE can be removed without further disruption (*18*). The PsaE subunit therefore appears necessary for proper association of ferredoxin or flavodoxin with PS I. A recent comparison of the wild-type and PsaE-less PS I complexes in *Synechocystis* sp. strain PCC 6803 also suggests that PsaE modulates the lifetime of the PS I/ferredoxin complex (*21*).

PsaE proteins are small (8−11 kDa) and highly basic. To date, complete PsaE sequences from 18 sources are available in the Swiss-Prot database. The solution structure of one of these (PsaE subunit from *Synechococcus* sp. strain PCC 7002, henceforth "S7002 PsaE") has been determined by NMR methods; it consists of a five-stranded antiparallel β-sheet with the topology of an SH3 domain (*22*). There is a high level of sequence similarity within the PsaE family, and the cyanobacterial PsaE proteins display nearly 70% identity (*23*). Two groups of PsaE proteins can be defined according to the length of the loop connecting the third β strand (or βC) and fourth β strand (or βD). The PsaE proteins of all eukaryotes, as well as those from the filamentous, heterocyst-forming cyanobacteria *Nostoc* sp. PCC 8009, *Calothrix* sp. PCC 7601, *Mastigocladus laminosus*, and *Anabaena variabilis* strain ATCC 29413, have a seven-residue deletion within this CD loop that is not observed in other prokaryotic PsaE proteins. PsaE proteins from higher plants, such as those from spinach and barley, also contain an N-terminal extension of about 20 residues relative to the cyanobacterial proteins. All PsaE proteins except that from *Synechococcus* sp. PCC 7002 have a sequence of residues close to either the N- or the C-terminus that is rich in alanine, proline, and lysine residues.

The crystal structure of the PS I complex of *Synechococcus elongatus* was determined to a resolution of 4 Å a few years ago (*2*, *4*). A number of transmembrane helices from PsaA and PsaB could be distinguished, and the general arrangement of the subunits established, but many details, including the position and orientation of PsaE on the membrane, remained unknown. An improved model was recently constructed with the aid of the S7002 PsaE NMR structure (*24*). In this model, the relative position and orientation of all three stromal proteins are visible. The *S. elongatus* PsaE protein makes contact with the membrane surface through the AB loop, and the large CD loop is embedded under PsaC, near the F_X cluster. Strand βB is perpendicular to and extends away from the surface, whereas strand βC is directed toward it. Both strands βA and βE are nearly parallel to the membrane surface and point the N- and C-termini into the stroma. This positioning is consistent with previous biophysical and structural studies of the stromal proteins, including cross-linking data.

The existence of two classes of PsaE proteins requires the determination of a second representative structure in order to anticipate the fold of related proteins and explore their interactions with PS I subunits. The consequences of the deletion in the CD loop and the C-terminal extension in the isolated PsaE subunit were examined in this work with the PsaE protein from PS I in *Nostoc* sp. strain PCC 8009 ("N8009 PsaE"). Its three-dimensional structure offers insight into the functional significance of the sequence changes and allows for modeling of PsaE proteins from other sources including higher plants.

MATERIALS AND METHODS

Protein Purification and Preparation. All reagents were purchased from Sigma (St. Louis, MO) unless otherwise indicated. *E. coli* strain BL21(DE3) cells were made competent according to the calcium chloride method (*25*) and transformed with pET3d plasmid containing the *psaE* gene from *Nostoc* sp. strain PCC 8009 (GenBank Accession Number AF148219) according to the Hanahan method (*25*, *26*). Plasmid stocks were prepared using either the Wizard miniprep (Promega, Madison, WI) or the Jetstar midiprep (Genomed, Research Triangle Park, NC).

Protein overexpression in M9 minimal media was achieved utilizing the T7 promoter (*27*) as described for *Synechococcus* sp. strain PCC 7002 (*13*). Inclusion bodies were solubilized by thorough resuspension in 4 mL of 5 × TS buffer (20 mM Tris-HCl, 10 mM NaCl, pH 7.8), and addition of 16 mL of 10 M urea. The solution was stirred at room temperature for 30 min, and insoluble material was removed by centrifugation at 48000*g* for 60 min at 4 °C. The supernatant, containing the solubilized denatured protein, was decanted and loaded directly onto a Sephadex G-75 gel filtration column equilibrated with TS buffer. The protein refolded on the column, and native PsaE eluted at a flow rate of ca. 0.35 mL min^{-1}.

Fractions containing PsaE [as ascertained by UV−vis spectroscopy and SDS−PAGE analysis (*28*)] were pooled and subjected to treatment with streptomycin sulfate to precipitate nucleic acid impurities. The supernatant was passed through the Sephadex G-75 column a second time. Purity was estimated at better than 95% according to SDS−PAGE analysis and Coomassie blue staining (*29*).

Uniformly labeled ^{15}N and $^{15}N/^{13}C$ N8009 PsaE was produced by growth of *E. coli* cells in M9 media supplemented with [^{15}N]ammonium chloride (1 g L^{-1}) or [^{15}N]-ammonium chloride and [^{13}C]glucose (2 g L^{-1}) (Isotec, Miamisburg, OH). Overepression and purification of the protein were carried out as described above.

To determine the extent of cleavage of the N-terminal methionine, a small sample of the PsaE protein was purified by reverse-phase HPLC on a Waters system (Milford, MA). The column (hydrophobic, C8) was loaded with 90% (v/v) acetonitrile and 0.1% trifluoroacetic acid (v/v) as solvent mixture; a gradient of 10−75% (v/v) trifluoroacetic acid was used for elution. The protein was lyophilized and redissolved in water to a final concentration of 100 pM. The protein was subjected to five rounds of Edman degradation (performed at the Hershey Medical Center, Hershey, PA). The N-terminal Met was found in ∼24% of the molecules. This result was confirmed by MALDI mass spectrometry (performed at the Wistar Institute Mass Spectrometry Facility), which identified two protein components $(M+H)^+ = 7961$, corresponding to PsaE starting at Met 1 (expected mass 7958 Da), and $(M+H)^+ = 7829$, corresponding to PsaE starting at Val 2 (expected mass 7827 Da)].

The extinction coefficient of the PsaE protein was determined by the method of Gill and von Hippel (*30*). Stock solutions of phosphate buffer and 8 M guanidine hydrochloride in phosphate buffer were prepared, as well as protein stock solutions at concentrations on the order of 15−40 mg mL^{-1}. The protein was added to each of the buffer and guanidine hydrochloride solutions and allowed to equilibrate for at least 1 h. Three samples were prepared at each protein concentration to determine reproducibility. Absorbance spectra were collected on an Aviv 14-DS spectrophotometer (Aviv Associates, Lakewood, NJ). The calculated value at 280 nm for the unfolded protein is 9.5 mM^{-1} cm^{-1}. The value determined experimentally as used in this work for the folded protein is 7.1 mM^{-1} cm^{-1}.

Protein Denaturation. Solutions with varying concentrations of urea were prepared by combining appropriate volumes of a 9 M urea stock solution (20 mM phosphate, pH 5.0) with phosphate buffer. A protein solution (20 mM phosphate buffer, pH 5) was added to each of the 30−45 samples so as to obtain a protein concentration of either 5 μM or 20 μM. The samples were left to equilibrate at room temperature for at least 1 h before data collection. The same samples were used for CD and fluorescence analyses. After measurement, the index of refraction of each sample was determined on an Abbe Mark II digital refractometer (Cambridge Instruments, Buffalo, NY) to determine the exact urea concentration (*31*). Reversibility was assessed by mixing 0 and 9 M samples together to create intermediate concentrations, which were then compared to the original samples. By this method, reversibility was better than 95%. Reversibility was inferior at pH 7.0.

CD data were collected on an Aviv model 62DS circular dichroism spectrophotometer (Aviv Associates). To obtain a native-state spectrum, data points were collected at 1 nm intervals from 320 to 220 nm with a 2 s averaging time at each wavelength. The wavelength (226 nm) at which the signal intensity was maximal in the native spectrum was chosen to monitor unfolding as a function of denaturant concentration. The signal at this wavelength was averaged for 120 s with sampling at a rate of 1 s^{-1}. The first 5 points were discarded, and the average of the remaining 115 points was used.

Fluorescence data were collected on either a Spex fluorolog 1.81m spectrophotometer or an Aviv model ATF 105 automated titrating differential/ratio spectrophotometer. The excitation wavelength was 295 nm, and emission was monitored from 305 to 460 nm. The unfolding curve was generated by extracting the intensity of the fluorescence at 352 nm from each spectrum and plotting it as a function of denaturant concentration. Data were also collected with excitation at 276 nm to observe tyrosine emission.

Thermal denaturation was carried out by following the change in the UV spectrum as a function of temperature. The 30 μM protein sample was prepared in 20 mM phosphate buffer at pH 7.0. Data were collected on an Aviv 14-DS spectrophotometer equipped with a thermoelectric device. The temperature was raised in steps of 2 °C; data were collected at 5 min intervals. Returning the sample to room temperature and comparing the spectra before and after cooling allowed for an assessment of the reversibility of folding, which was routinely better than 95%. The wavelength (295 nm) of maximal change was used to obtain denaturation curves, which were analyzed with the equations for a two-state unfolding process (*32*, *33*). The equations describing the chemical and thermal denaturations are provided in the Supporting Information.

NMR Spectroscopy. Lyophilized protein was dissolved in 0.56 mL of H_2O or ^2H_2O to obtain a final protein concentration of 1−2 mM. The pH or pH* (the pH uncorrected for isotope effects) of the sample was adjusted to 6.1 using ^2HCl or NaO^2H, to give a final ^2H_2O concentration of 5−10% of the total H_2O sample volume.

Data were collected on a Bruker AMX2 spectrometer at 500.13 MHz (^1H), 125.76 MHz (^{13}C), and 50.68 MHz (^{15}N). Sample temperature was 298 K (used for structure determination) or 290 K to resolve accidental chemical shift degeneracies. One-dimensional data were also collected up to 323 K. Solvent suppression was achieved via a 1.2 s low-power presaturation pulse, by utilizing pulsed field gradients (*34*), or through the use of a water flip-back-WATERGATE scheme. The ^1H spectra were referenced to the water resonance at 4.76 ppm; ^{13}C signals and ^{15}N signals were indirectly referenced to the proton frequency (*35*).

Homonuclear proton spectra (2QF-COSY, 2Q, TOCSY, and NOESY experiments) were collected with standard 2D pulse sequences as described previously in both 95% H_2O/5% ^2H_2O and 99.96% ^2H_2O (*36*). NOESY data were collected with mixing times of 100 and 70 ms; the latter was used for cross-peak volume measurements in the structure calculations.

A series of $^1H-^{15}N$ HSQC data sets (*37*) were collected to determine the rate of exchange of backbone amide protons against deuterons. A sample of lyophilized protein was dissolved in ^2H_2O (pH 6.1), and data acquisition was initiated within 1.5 min of dissolution. After 24 h of exchange, all signals were lost. The intensity of 27 amide cross-peaks could be followed with time and fit to a simple exponential decay function to obtain first-order exchange rate constants, k_{ex}. Intrinsic exchange rate constants (k_{int}) were calculated for each of these amides by taking into account the primary structure of the protein, temperature, and pH as described

by Bai and co-workers (*38*). If a cross-peak remained after 3 h in the $^1H-^{15}N$ HSQC spectra, a hydrogen bonding partner was sought in the structure. H-bonds recognized in a majority of early structures were implemented as restraints.

A series of nine *J*-modulated $^1H-^{15}N$ HSQC spectra (*39*, *40*) were collected to determine $^3J_{\alpha H-NH}$ coupling constants and from these, backbone φ dihedral angles, as described for S7002 PsaE (*36*). Stereospecific assignments for β-methylenes, valine, and leucine methyls, and glycine α-hydrogens were made as discussed previously (*22*).

The following 3D experiments were run with their indicated spectral widths and complex data points (*) acquired for dimensions t_1, t_2, and t_3, respectively: $^1H-^{15}N$ TOCSY−HSQC (^1H: 6024 Hz, 64*; ^{15}N: 1700 Hz, 32*; ^1H: 6024 Hz, 1024*) (modified from ref *41*); $^1H-^{15}N$ NOESY−HSQC (^1H: 6024 Hz, 58*; ^{15}N: 1700 Hz, 32*; ^1H: 6024 Hz, 1024*) (modified from ref *41*); $^1H-^{13}C$ HCCH−TOCSY and COSY (^1H: 4000 Hz, 64*; ^{13}C: 8000 Hz, 32*; ^1H: 4000 Hz, 512*) (*42*); $^1H-^{13}C$ NOESY−HSQC (^1H: 6024 Hz, 64*; ^{13}C: 7575 Hz, 32*; ^1H: 6024 Hz, 512*) (*43*); HNCA (^{15}N: 1700 Hz, 36*; ^{13}C: 2500 Hz, 48*; ^1H: 6024 Hz, 512*) (*44*, *45*); CBCANH (^{13}C: 7142 Hz, 50*; ^{15}N: 1700 Hz, 36*; ^1H: 7024 Hz, 512*) (*46*); and CBCA(CO)NH (^{13}C: 7142 Hz, 50*; ^{15}N: 1700 Hz, 36*; ^1H: 7024 Hz, 512*) (modified from ref *47*). In all these experiments, quadrature detection in the indirect dimensions was achieved using the TPPI−States method (*48*).

Structure Calculations. Structures were calculated using the program X-PLOR 3.851 (*49*) with the protocol based on a hybrid distance geometry−simulated annealing (DG-SA) method (*50*). A total of 1093 restraints were applied as detailed under Results and Table 1. The iterative approach employed for the first generation of S7002 PsaE structures (*22*) was applied with the following modifications. DG-SA structures (a total of 100) were calculated from embedded substructures and sorted by X-PLOR energy. These structures were then subjected to an SA-refinement that consisted of 4000 slow-cooling steps from the initial temperature of 1000 K to 100 K. Finally, the 30 structures with lowest X-PLOR energy were subjected to 8000 steps of restrained minimization to provide the final accepted models. A family of the 20 lowest energy structures without significant distance violations (all $\leq$0.02 Å as reported by X-PLOR) is presented in this work.

Structures and restraints were analyzed with the programs AQUA and PROCHECK (*51*). The checking facility of the program WHAT IF (*52*, *53*) was used to compare the calculated structures with a set of high-resolution protein structures and to identify unusual geometries in the backbone and side chains.

RESULTS

The N8009 and S7002 PsaE proteins are 69% identical and 83% similar in amino acid sequence (Figure 1). As notable differences, S7002 PsaE possesses a seven-residue insertion in the CD loop, and N8009 PsaE has an eight-residue C-terminal extension of alanines, prolines, and lysines. The only replacements in the 22 residues of the β-sheet occur at positions 9 (Lys in S7002 PsaE to Arg in N8009 PsaE), 66 (Glu to Val), and 67 (Val to Glu).

The ^1H NMR spectrum of N8009 PsaE is typical of that for a small, well-folded protein and comparable to that of

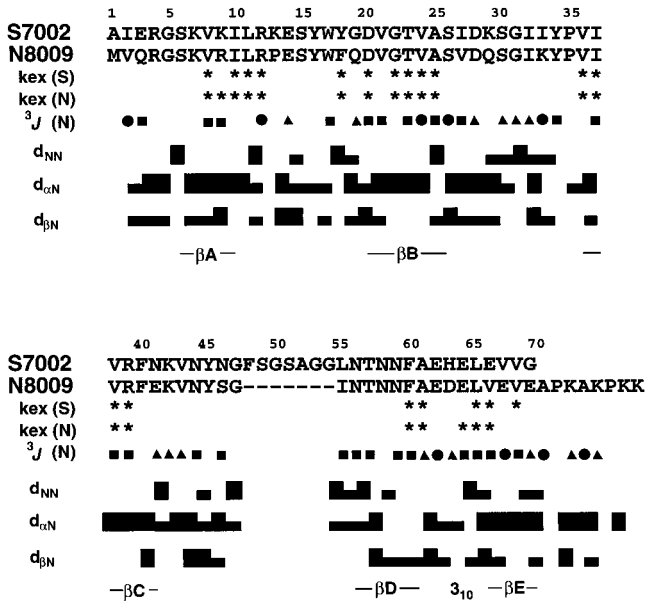

FIGURE 1: Primary structures for the N8009 and S7002 PsaE proteins. The numbering of the N8009 PsaE protein is Y45-S46-G47-I55 for consistency with the S7002 PsaE numbering scheme and in accordance with structural arguments (see text). The lines marked kex (S) and kex (N) indicate the slowly exchanging backbone amide protons in S7002 and N8009 PsaE, respectively. The exchange is defined as slow if the signal remains detectable after a 5 h exposure to ^2H_2O at pH ~6 and 298 K. The line marked ^3J (N) contains the NHCαH coupling constants: (■) >8 Hz; (▲) between 6 and 8 Hz; (●) <6 Hz. Sequential NOE information for the N8009 PsaE protein at 298 K and pH 6.1 is shown in the next three lines, with the height of the bar corresponding to medium (short bar) or strong (tall bar) effects; d_{NN} represents the NH_i-NH_{i+1} contacts, $d_{\alpha N}$ represents the $C\alpha H_i-NH_{i+1}$ contacts, and $d_{\beta N}$ is the $C\beta H_i-NH_{i+1}$ contacts. The last line contains the secondary structure assignments as determined by the program DSSP (*78*) on the final structures.

S7002 PsaE (*36*). For N8009 PsaE and S7002 PsaE alike, chemical shift degeneracy rendered heteronuclear data necessary to bring to near completion the proton resonance assignment step. The ^1H, ^{13}C, and ^{15}N chemical shifts obtained with standard 2D and 3D experiments are listed in Table S1 of the Supporting Information. Few distinctive spectral features of N8009 PsaE were noticed in the early stages of analysis. The most striking involves the side chain of Phe40. In S7002 PsaE, Phe40 has readily detectable δ and ϵ ring ^1H signals at room temperature. In N8009 PsaE, these ^1H signals are broad and require temperatures higher than 300 K for unambiguous identification. This behavior is suggestive of altered dynamic properties in the βC region of the structure.

Two other unique characteristics of N8009 PsaE are apparent in the $^1H-^{15}N$ HSQC data shown in Figure 2, which display doubling of resonances in the N- and C-terminal regions of the protein. At the N-terminus, the heterogeneity is chemical and results from partial cleavage of the terminal methionine. In the C-terminal region, the heterogeneity is conformational and attributed to cis-trans isomerization of the Ala70−Pro71 linkage. The affected backbone NH signals are those for residues (66−70) near and within βE and residues (6−8) on the adjacent strand in βA. The most intense peak in the pair, when resolved, was used for calculating the structures.

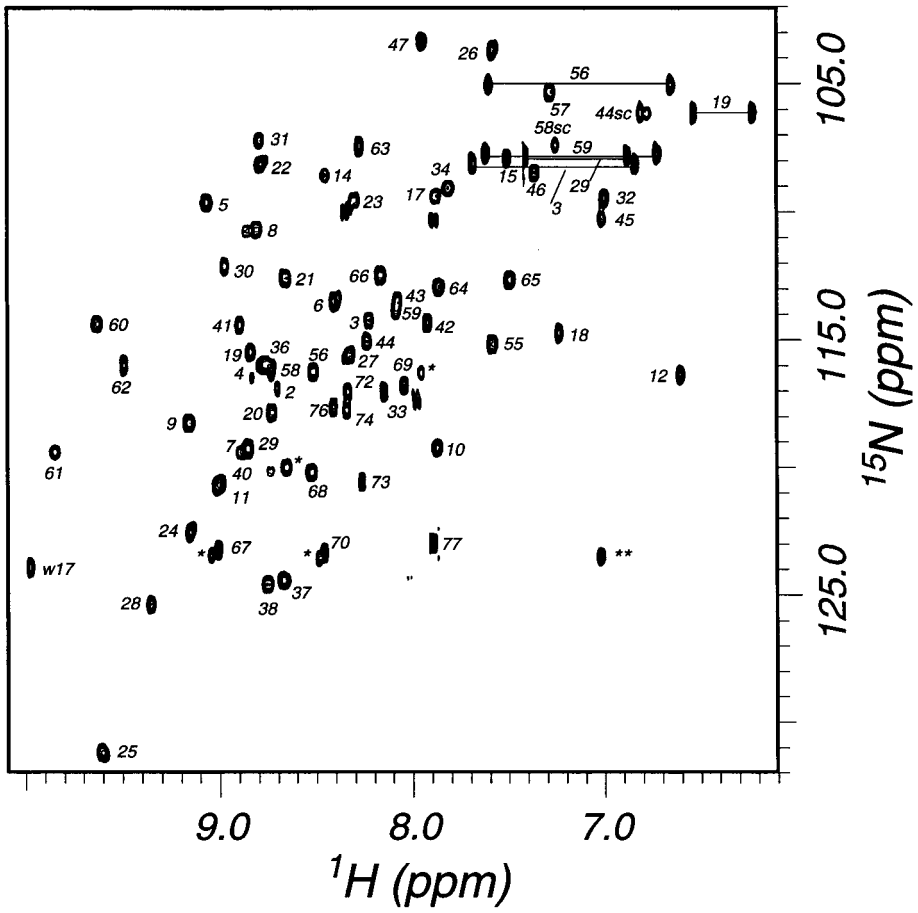

FIGURE 2: ^{1}H−^{15}N HSQC spectrum of N8009 PsaE at 298 K, pH 5.8. Assignments for the amide backbone resonances are indicated as well as the indole NH of Trp17 (marked w17). The amide groups of Asn side chains are indicated with horizontal lines or the letters "sc". A folded Arg NϵH cross-peak is indicated with a double asterisk, and minor second forms of specific residues are indicated with an asterisk. These minor forms arise from an apparent Xxx-Pro cis/trans isomerization for residues 70−71. The effects of this isomerization are seen in βE and βA. Other minor peaks arise from partial cleavage of the terminal Met.

To ascertain the integrity of the β-structure in N8009 PsaE, a comparison of the rate at which backbone amide hydrogens exchange with solvent was undertaken. At pH 6.1 (uncorrected for isotope effects) and 298 K, the backbone NH protection factors (k_{int}/k_{ex}, see Materials and Methods) adopt values on the order of 10^4. Differences with S7002 PsaE (*36*) are an enhanced protection for the backbone NH of Arg9 and Glu64 in N8009 PsaE, and an enhanced protection of Val68 in S7002 PsaE. Otherwise, the pattern of hydrogen-bonded (or buried) amide protons is identical. The location of the slowly exchanging protons in both proteins is indicated in Figure 1.

The NOESY data provided over 960 distance restraints, mostly concentrated over residues 2−68. To these were added H-bond restraints and dihedral angle restraints (backbone φ and side chain χ angles) as detailed in Table 1. On average, there are approximately 16 restraints per folded residue. The X-PLOR structural calculations generated a family of acceptable conformations; the 20 structures with the lowest X-PLOR energy are superimposed in Figure 3. Figure 4 contains the ribbon representation of the average structure. The rmsd over the regular secondary structure is 0.24 Å for backbone atoms, and the rmsd for interior side chains is 0.63 Å. These and other structural statistics are summarized in Table 1. The rmsd values indicate that the backbone atoms involved in secondary structure and most nonsurface side chains are well constrained by the data. According to the Ramachandran plot, the equivalent resolution of this structure is 2.5 Å (*51*), if the disordered C-terminal tail is excluded from the calculation. The structure evaluation tools of the program WHAT IF reveal the flaws usually found in X-PLOR NMR structures (in particular, modest Ramachandran quality and tight peptide planarity). A few extraordinary features and anomalies can be explained by the lack of constraints in loop regions and for exposed side chains.

Urea and Thermal Denaturation. Despite the negligible amino acid sequence identity displayed by PsaE proteins and SH3 domains, a strong structural similarity is observed between these apparently unrelated proteins. To compare the thermodynamic properties of the fold as manifested in different sequences, a denaturation study was undertaken with N8009 PsaE. The CD spectrum of the typical SH3 domain is complex (*54*−*56*), and so is the CD spectrum of PsaE. Native N8009 PsaE presents a trough of negative ellipticity around 243 nm, a peak of positive ellipticity at 225 nm, and another region of negative ellipticity below 217 nm. The spectrum undergoes a significant change upon the addition of urea, becoming negative at wavelengths below 250 nm. A wavelength of 226 nm was chosen to monitor the denaturation. Typical raw data are shown in Figure 5A.

Table 1: Structural Statistics and Atomic Rms Deviations

Structural Statistics	{SA}[a]	⟨SA⟩$_r$[b]
rms deviation from experimental constraints	0.0007 ± 0.0001	
distance constraints (Å)[c]	0.0007 ± 0.0001	0.00085
dihedral constraints (deg)[d]	0.04 ± 0.03	0.029
rms deviation from ideal geometry		
bonds (Å)	0.00252 ± 0.00003	0.0025
angles (deg)	0.526 ± 0.001	0.53
impropers (deg)	0.77 ± 0.06	0.82
number of violations[e]		
distance constraints	37 ± 4	30
dihedral constraints	9 ± 1	10

Rms Deviations from the Average Structure (Å)[f]	all residues[g]	limited set[h]	regular 2° structure[i]
backbone (N−Cα−C)	0.52 ± 0.12	0.30 ± 0.05	0.24 ± 0.04
side chains (heavy atom)	1.07 ± 0.12	0.95 ± 0.11	0.65 ± 0.07

[a] {SA} is the family of final 20 simulated annealing structures; mean values and standard deviations are shown for these 20 structures. [b] ⟨SA⟩$_r$ is the structure obtained by applying 10 000 steps of restrained minimization to the mathematical average structure. The mathematical average structure was obtained by first aligning the 20 structures with a least-squares fit over the regular secondary structure [(N−Cα−C) of residues 6−10, 21−26, 36−39, and 57−68] and then averaging the atomic coordinates. [c] A total of 1020 distance constraints were used in the structure calculation: 303 were long range; ($|i - j| > 4$); 305 were medium range ($1 < |i - j| \leq 4$); 353 were intraresidue; and 27 * 2 were H-bonds. [d] A total of 73 dihedral constraints were used in the structure calculation: 41 were backbone φ angles; 32 were side chain χ_1 angles. [e] *All* distance violations are less than 0.02 Å, and *all* dihedral violations are less than 1.3°. [f] The rms deviations of the atomic positions were calculated for each structure in {SA} against the mathematical average structure. The average values and standard deviations for the 20 final structures are reported. [g] Rmsd for residues 4−68. [h] Rmsd for residues 4−27, 34−40, and 57−68. This set excludes the disordered BC and CD loops and the C- and N-terminal residues. [i] Rmsd of the segment adopting regular secondary structure for backbone atoms (residues 6−10, 21−26, 36−39, and 57−68) and all nonsurface hydrophobic side chain atoms (8, 10, 13, 16−18, 21−25, 34−38, 40, 43, 55, 60, 61, 66 and 68).

Similar data are obtained by following the fluorescence of the protein. The maximum fluorescence emission wavelength of native N8009 PsaE protein is ∼352 nm and is due to the lone tryptophan, Trp17. When the protein is excited at 276 nm, there is no signal in the native spectrum corresponding to emission from the three tyrosine residues, presumably because of efficient energy transfer to the tryptophan residue. Upon unfolding, the tryptophan emission spectrum undergoes a red shift of approximately 12 nm, with a concomitant decrease in intensity. This is typical of increased solvent accessibility upon denaturation (*57*). As the protein unfolds, a tyrosine emission signal becomes visible at 300 nm, indicating that some of the tyrosines have moved sufficiently far away from the tryptophan for energy transfer to be less efficient. Raw fluorescence data are shown in Figure 5B.

The thermodynamic parameters obtained through monitoring the fluorescence changes at 352 nm are within error of those extracted from the CD data and support the two-state model of denaturation. Figure 5C shows the apparent fraction of unfolded protein calculated with a two-state model of denaturation. In Figures 5A and 5B, the solid line is the result of the fit and supports that only two states are detectable by this method. When the CD and fluorescence data are analyzed with a global fitting procedure, the values of ΔG°_{H2O}, m, and $[\text{urea}]_{1/2}$ for N8009 PsaE are 12.4 ± 0.3 kJ mol^{-1}, 3.6 ± 0.1 kJ mol^{-1} M^{-1}, and 3.49 ± 0.02 M, respectively.

A thermal denaturation experiment was carried out at pH 7.0, by monitoring absorption at 295 nm (data not shown). The transition is reversible, and data can again be fit satisfactorily to a two-state model, yielding a midpoint denaturation temperature of 59 °C. The enthalpy of denaturation is on the order of 225 kJ mol^{-1} at this temperature. The small change in heat capacity upon denaturation cannot be determined accurately with this method.

DISCUSSION

Structure of N8009 PsaE. A qualitative assessment based on primary structures predicts N8009 PsaE to have the same global fold as S7002 PsaE. The NMR data confirm that the five strands of the antiparallel β-sheet are comprised of residues 7−10 (βA), 21−27 (βB), 36−39 (βC), 57−60 (βD), and 65−67 (βE). The 3_{10} helical turn between βD and βE (residues 62−64) is apparent through CβH$_i$−NH$_{i+3}$ NOEs and elevated protection factors for Glu64 and Leu65. The AB loop, which connects strands βA and βB, is formed by residues 11−19 and is the best defined of the three loops. The BC loop, joining strands βB and βC, has a lower average number of NOE restraints per residue than the rest of the protein. As a result, several distinct conformations are compatible with the experimental data, increasing the local backbone rmsd value. The CD loop displays a degree of conformational heterogeneity comparable to that of the BC loop. Finally, the eight residues of the C-terminal tail do not show any dipolar interactions with the rest of the protein. The backbone amide resonances of this segment are sharp; they exhibit low protection factors, and their chemical shifts have a steep slope with respect to temperature (∼8 ppb K^{-1}). These spectral indicators support the interpretation that the conformational freedom of the tail is limited only by residues Pro71 and Pro75.

N8009 PsaE has a small hydrophobic core consisting of Val8 (βA), Ile10 (βA), Val24 (βB), Val36 (βC), Val38 (βC), Phe60 (βD), and Leu65 (βE). Phe40 (CD) stacks at the CD edge of the barrel and closes its entrance; Trp17 (AB) lies parallel to this residue, points toward the outside of the protein, and has limited solvent accessibility as well. The NϵH of Trp17 is within hydrogen-bonding distance of the backbone oxygen of Lys42. It displays increased protection to exchange consistent with formation of the interaction. In addition, two short-range, intramolecular salt bridges are likely: between Arg12 and Glu14, and between Arg39 and Glu41. Arg12 shows exchange-broadened side-chain resonances that might reflect interactions with Glu14 and Tyr34. Arg39−Glu41 interactions, if present, are NMR-silent in the current data.

Among the PsaE proteins listed in the Swiss-Prot database, *Nostoc* sp. strain PCC 8009 PsaE is closest in sequence space to *Mastigocladus laminosus* (*Fischerella* sp.) PsaE (86% identity, 97% similarity), *Fremyella diplosiphon* (*Calothrix* sp. strain PCC 7601) PsaE (89% identity, 94% similarity), and *Anabaena variabilis* PsaE (91% identity, 95% similarity). The PsaE proteins of *Cyanophora paradoxa* PsaE (78%

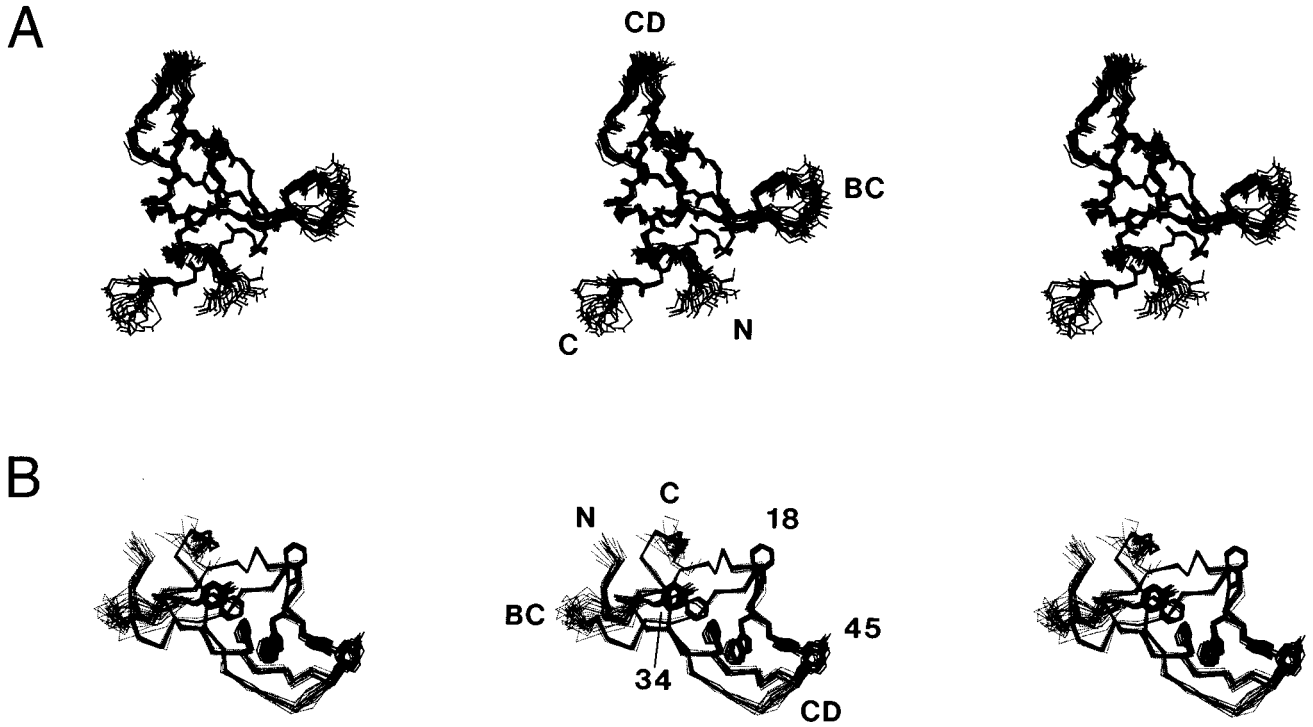

FIGURE 3: Stereo triptych of 20 superimposed structures of N8009 PsaE calculated as described under Materials and Methods. The backbone atoms in the regular secondary structure were used for this alignment. (A) The backbone from residue 1 to residue 70 (the disordered C-terminal tail has been deleted). The orientation is the same as in Figure 4. (B) A different view showing the Cα trace and the aromatic residues. The CD loop points to the lower right. Phe40 is the residue nearly parallel to Trp17 recognizable in the center.

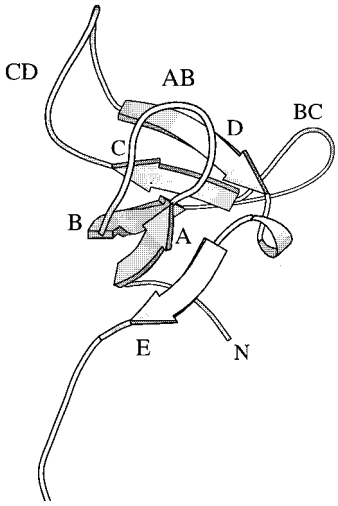

FIGURE 4: Ribbon structure of the lowest energy structure of the N8009 PsaE protein. Like the S7002 protein, it is a five-stranded β-sheet with a 3_{10} helical turn between the penultimate and last strands. The C-terminal tail is disordered.

identity, 93% similarity), *Cyanidium caldarium* (*Galdieria sulfuraria*) PsaE (71% identity, 93% similarity), *Guillardia theta* (*Cryptomonas phi*) PsaE (77% identity, 95% similarity), and *Odontella sinensis* PsaE (67% identity, 91% similarity) are also close. The structure presented here can serve as an excellent model for all of these bacterial and eukaryotic proteins. Interestingly, *M. laminosus* and *C. caldarium* are moderately thermophilic organisms. It is unclear which, if any, of the few amino acid replacements confer thermostability.

Comparison of S7002 and N8009 PsaE Proteins. The refined structures of the two PsaE proteins (this work; Falzone et al., unpublished data) are superimposed in Figure 6, which illustrates the high similarity of the regular secondary structure. The backbone heavy-atom rmsd between the average structures in these regions is 0.6 Å. Spectroscopically, differences are detected in the weakest interstrand NOEs, for example, that between Val24 NH and Lys7 αH, seen in S7002 PsaE but not in N8009 PsaE. These and other minor NOE discrepancies do not translate into significant deviations in the backbone of the two proteins. Divergences between the two structures are seen in the connecting loops, perhaps as the result of a low number of restraints. A real variance is noted for the buried Phe40 whose broad ring resonances are indicative of further mobility restriction at 298 K in N8009 compared to S7002. Interestingly, the residues making up the hydrophobic core are identically conserved. Thus, the cause of the restricted rotation in N8009 could be the collapsed and more rigid CD loop hindering the motions of Trp17 and therefore Phe40.

In the AB region, the aromatic cluster Tyr16-Trp17-Phe18 (Tyr18 in S7002 PsaE) is partially exposed to solvent. Complete sets of NOEs are observed in both proteins from these aromatic side chains to other hydrophobic groups, including Phe40. These residues participate in an extensive aromatic network that is particularly visible in the N8009 PsaE structure (Figure 3B). Starting from one side of the protein and extending to the other, the cluster includes Tyr45, Tyr16, Trp17, Phe40, Phe60, and Tyr34. It remains to be elucidated whether this striking structural formation plays a special role in electron transport in vivo.

The BC loop is similarly disordered in both proteins. Only one conservative substitution (Ile27 in S7002 PsaE to Val in N8009 PsaE) and one nonconservative substitution (Ile33 in S7002 PsaE to Lys in N8009 PsaE) are present. The latter

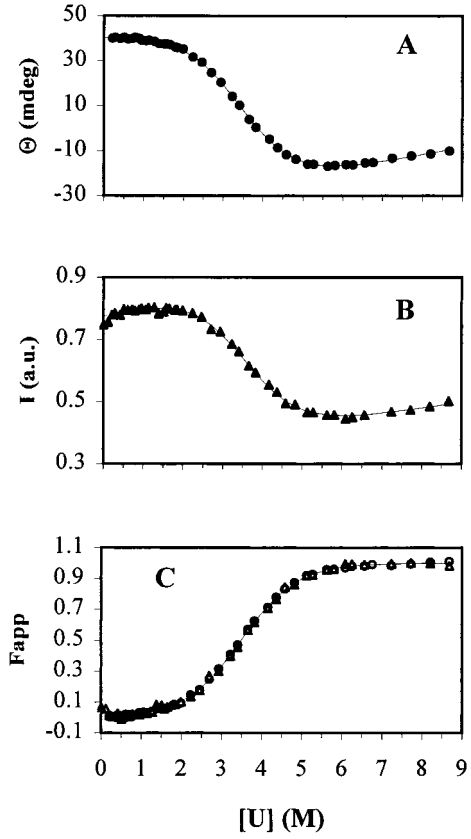

FIGURE 5: Plot of the CD signal at 226 nm (A) and the fluorescence emission at 352 nm (B) for N8009 PsaE as a function of urea concentration. The protein concentration is 20 μM in 20 mM phosphate buffer, pH 5. The data were collected at 298 K on the same samples. The solid lines through the data points were calculated with the parameters of the global fit of the CD data and the fluorescence data collected at this concentration and at 5 μM protein. (C) Apparent fraction of unfolded protein versus urea concentration derived from the CD data (○) and the fluorescence data (△) shown in panels A and B. The solid line was calculated with the parameters of the global fit of the CD data and the fluorescence data collected at 20 and 5 μM protein.

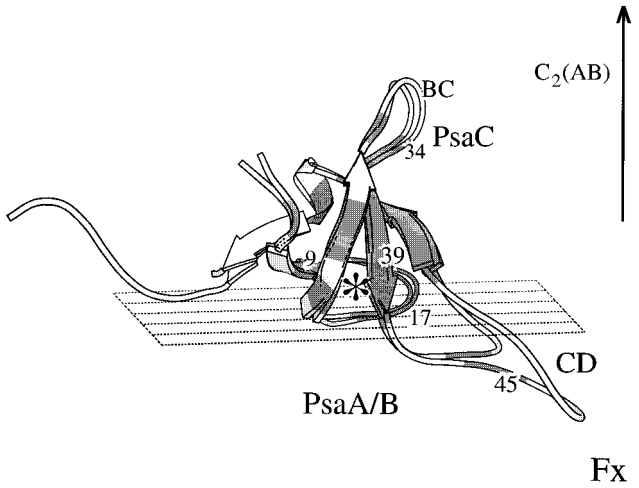

FIGURE 6: Superimposed ribbon structures of the S7002 and N8009 PsaE proteins. The orientation is similar to that in Figure 5 of Klukas and co-workers (*24*). The location of conserved or rarely substituted residues is marked in gray. The approximate location of the membrane surface is shown by a horizontal plane. PsaE contacts PsaA/B through the AB loop and the CD loop. The tip of the CD loop points toward the F_X cluster. The BC loop interacts with the central loop of PsaC. The axis marked C_2(AB) indicates the local PsaA/B axis of symmetry. The star indicates the end of βB and the beginning of βC, where ferredoxin docks onto PsaE.

replacement breaks the hydrophobic stretch extending from residue 32 to residue 38 in S7002 PsaE. In S7002 PsaE, the large CD loop has increased mobility relative to the rest of the protein (*22*), and its structure is underconstrained by the NMR data. In the N8009 PsaE protein, the short CD loop imposes negligible rearrangement at the tethered ends of the βC and βD strands.

The three-dimensional overlap shown in Figure 6 suggests a structure-based sequence alignment of the two PsaE subunits. The Asn residue preceding Thr57 in N8009 PsaE is the structural equivalent of Asn56 in S7002 PsaE. It has the same αH−αH NOE to Val43 and the same side chain orientation. On the other hand, the Ser residue following Tyr45 in N8009 PsaE has a close match in Asn46. Thus, the deletion can be placed from residue 48 to residue 54. This involves two plausible replacements at or near the edges: Asn46 to Ser and Leu55 to Ile. The sequence alignment shown in Figure 1 was derived from this structural information.

The N8009 PsaE CD loop exhibits long-range contacts involving Tyr45 and Tyr16 that are not observed in S7002, even though in this protein, the NOEs between the adjacent residues Asn44 and Val43 and Trp17 are observed. It is possible that in S7002 PsaE, the mobile CD loop rarely visits the conformation yielding the NOEs between residue 16 and residue 45. Interestingly, the ellipticity at 225 nm is reduced by a factor of nearly 4 in S7002 PsaE compared to N8009 PsaE (*29*). Tertiary interactions between the aromatic side chains can contribute significantly to the CD signal around 225 nm in proteins with low helix content (*57*). Given that the secondary structures of the PsaE proteins are the same and that the spatial relationship of Trp17 with respect to other aromatic side chains is largely invariant, the spectral discrepancy is likely to reflect the altered residue 16 to residue 45 interactions. Finally, in the CD loop region corresponding to the deletion, Phe48 of S7002 PsaE has a weak NOE to Tyr45. This medium-range effect apparently results from a preferred hydrophobic interaction in solution.

Stability of the PsaE Fold. The structural similarity with the SH3 domain raises the possibility that other properties of the fold, such as thermodynamic stability, are maintained. The denaturation of N8009 PsaE by urea can be modeled as a simple equilibrium between a native state and an unfolded state. Two-state behavior is also observed for SH3 domains, including those of Sem (*58*), spectrin (*54*, *59*), drk (*60*), Sso7d (*56*), Src (*61*), Btk (*55*, *62*), Itk (*62*), Tec (*62*), Fyn (*63*), and P13 (*64*). The range of $\Delta G°$ values obtained for these proteins extends from 11 to 17 kJ mol^{-1}, with the exceptions of Fyn (25 kJ mol^{-1}) and Sso7d (33 kJ mol^{-1}). The PsaE protein lies at the lower end of the range. Similarly, the dependence of $\Delta G°$ on urea concentration (or *m* value), which measures the change in exposed hydrophobic surface area upon unfolding, is 3.3 kJ mol^{-1} M^{-1} in N8009 PsaE. This value is also at the low end of the range covered by the SH3 domains mentioned above (2.9−9.6 kJ mol^{-1} M^{-1}). Many of these exhibit a T_m above 70 °C (*55*, *58*, *62*), contrasting with their low room-temperature stability. However, a high T_m does not accompany the low stability of N8009 PsaE with respect to chemical denaturation, and the

onset of unfolding occurs near 40 °C at pH 7.0. The thermodynamic stability of the isolated protein, lower than expected for a protein of this size (*65*), may explain the small level of PsaE detected in the cell when the complex is not assembled (*66*). It is also possible that additional stability is provided to the PsaE fold by the complex formation with the adjacent PS I components.

Function and Location of PsaE Proteins. The structure of the N8009 protein and the improved model of PS I (*24*) provide insight into the potential roles of specific PsaE amino acids. The strictly conserved, nearly conserved, and variable residues of the PsaE protein are highlighted in Figure 6. The orientation of the ribbon corresponds to that presented by Klukas and co-workers [Figure 5 in (*24*)], with the membrane plane lying horizontally under the protein.

The stretch Glu-Ser-Tyr-Trp (14−17) of the AB loop is strictly conserved in all known sequences, and residue 18 is usually a Phe or a Tyr. The AB loop is in contact with PsaA/B in the thylakoid membrane, which may explain the low occurrence of changes in this region. Interestingly, the Trp17Ala replacement in *Synechocystis* sp. strain PCC 6803 PsaE, whose sequence is closely related to that of *Synechococcus* sp. strain PCC 7002 PsaE, has no apparent consequence on the binding of PsaE to the PS I complex or on the reduction of ferredoxin (*17*). Likewise, in the N-terminal region, the conserved Arg12 can be replaced individually with an alanine without effect. In combination with replacements at Lys7 and Lys11, or Arg4 and Arg9, this change interferes with complex assembly. This suggests that single replacements can be compensated by the interprotein interactions, whereas multiple replacements may destabilize the protein and its complex excessively.

Other residues are involved in intermolecular contacts and are strictly conserved. Among these, Tyr34 is in the BC loop and is proximal to PsaC, and Arg39 is at the C-terminal end of βC, at the surface of the binding site for soluble ferredoxin. Ferredoxin must also bind proximally to PsaC in order to accept electrons from the F_B [4Fe-4S] cluster (*67*). PsaE has been shown to enhance the interaction between PsaC and ferredoxin (*17*, *68*, *69*). Replacement of Arg39 with glutamine decreases the affinity of PS I for ferredoxin by a factor of 20 (*70*, *71*). The PsaE structure reveals that the semiconserved Thr23 is on the surface of βB and proximal to Arg39; its replacement with a bulky residue would be expected to affect the interactions with ferredoxin.

Among the other conserved residues on the surface of the isolated subunit, there is an interesting trio consisting of Val43 (conserved in 17 of 18 PsaE sequences), Asn44 (conserved in 17 of 18 PsaE sequences), and Tyr45 (completely conserved). These residues make up part of the CD loop. As this loop is uncharged, it is not surprising to find it embedded between PsaC and PsaA/B, whose *e* and *f* helices and *n*−*o* loop it contacts (*24*, *72*). The N8009 PsaE CD loop contains seven residues fewer than S7002 PsaE or *S. elongatus* PsaE CD loops; however, in species that possess the CD loop deletion and that are closely related to *Nostoc* (e.g., *M. laminosus*), there are no amino acid insertions in the corresponding regions of PsaA/B that are proximal to PsaE. Since the PsaE subunit itself does not undergo major changes, it would appear that other subunits of PS I must rearrange to fill the cavity. It should also be noted that the N8009 and S7002 structures do not overlap perfectly in this region. It is likely that the conformations become more similar in the flexible loop regions when the subunit is part of the complex.

The PsaE subunit participates in multiple protein−protein interactions, as SH3 domains do. These domains bind a proline-rich polypeptide near the 3_{10} helix (*73*). The corresponding face in PsaE is in contact with PsaD and PsaC. There is no significant sequence similarity between the PsaE proteins and the SH3 domains, and the details of the PsaE contacts within the assembled PS I complex remain unclear. In addition, PsaE utilizes other elements of its structure to interact with PS I subunits, for example, the AB loop, the BC loop, and the CD loop. The latter two are both shorter in typical SH3 structures.

The C-terminal region of PsaE in *Synechocystis* sp. PCC 6803 has been implicated in the correct association of PsaC into the PS I complex (*70*). However, the C-terminus of PsaE points away from PsaC and the 3-fold symmetry axis and toward the exterior of the PS I trimer/monomer. In addition, this region of PsaE is accessible to proteases and susceptible to chemical modifications (*74*), and in *S. elongatus*, PsaE can be biotinylated at all seven of its lysine residues (three of which are in the C-terminal extension) as well as at its N-terminus (*75*). The lack of structure observed in the C-terminal tail of N8009 PsaE is consistent with solvent exposure in the complex.

Many of the cyanobacterial sequences, including Nostoc 8009, have a C-terminal extension with a net positive charge whereas higher plant PsaEs have negatively charged extensions of about 30 residues at the N-terminus (*76*). The orientation of the N- and C-termini explains how these terminal extensions can be readily accommodated in the structure. The portion of the PsaE sequences common to all organisms is that which interacts with PsaC, and, presumably, ferredoxin. The C- and N-termini appear to interact with the PsaF/J/M subunits at the exterior surface of the trimer and would not interfere with the packing of any other subunits on the complex.

In conclusion, the structure of *Nostoc* sp. strain PCC 8009 PsaE facilitates the interpretation of the biochemical and structural data gathered on PS I. The subunit demonstrates the versatility of the SH3 domain fold in mediating protein−protein interactions; PsaE binds to proteins on the stromal interface and, as the current structure shows, does so with a cluster of conserved and semiconserved residues on its surface. As PsaE is required for cyclic electron transport, it is likely that an additional role for these conserved residues is to maintain the proper assembly of proteins that participate in this process or are directly involved in this pathway around PS I, or both. N8009 PsaE also points to the plasticity of the stromal organization in the cyanobacterial reaction center as the void created by the smaller and potentially more structured CD loop is tolerated and somehow compensated in the complex. As a more detailed picture of PS I emerges, it will be feasible to determine the exact nature of the rearranged structure around the shorter version of the CD loop.

ACKNOWLEDGMENT

We are grateful to the members of Dr. C. R. Matthews's research group for help with the denaturation experiments

and to Drs. Krauss, Saenger, Klukas, and Schubert for communication of results prior to publication. We also thank Anne Stanley for the N-terminal analysis, Dr. Kaye Speicher for the mass spectrometry data, and WanLi Kuang for protein preparation. Figures 4 and 6 were generated with MOLSCRIPT (*77*), and Figure 3 with WHAT IF (*52*).

SUPPORTING INFORMATION AVAILABLE

A table summarizing the chemical shifts of the *Nostoc* sp. strain PCC 8009 PsaE; the details of denaturation data analysis; the fluorescence and CD spectra of native and unfolded *Nostoc* sp. strain PCC 8009 PsaE (7 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fromme, P. (1996) *Curr. Opin. Struct. Biol. 6*, 473−484.
2. Krauss, N., Schubert, W.-D., Klukas, O., Fromme, P., Witt, H. T., and Saenger, W. (1996) *Nat. Struct. Biol. 3*, 965−973.
3. Lelong, C., Boekema, E. J., Kruip, J., Bottin, H., Rogner, M., and Sétif, P. (1996) *EMBO J. 15*, 2160−2168.
4. Schubert, W.-D., Klukas, O., Krauss, N., Saenger, W., Fromme, P., and Witt, H. T. (1997) *J. Mol. Biol. 272*, 741−769.
5. Schubert, W.-D., Klukas, O., Saenger, W., Witt, H. T., Fromme, P., and Krauss, N. (1998) *J. Mol. Biol. 280*, 297−314.
6. Mi, H., Endo, T., Schreiber, U., Ogawa, T., and Asada, K. (1992) *Plant Cell Physiol. 33*, 1233−1237.
7. Böttcher, B., Gräber, P., and Boekema, E. J. (1992) *Biochim. Biophys. Acta 1100*, 125−136.
8. Kruip, J., Boekema, E. J., Bald, D., Boonstra, A. F., and Rögner, M. (1993) *J. Biol. Chem. 268*, 23353−23360.
9. Chitnis, P. (1996) *Plant Physiol. 111*, 661−669.
10. Golbeck, J. H., Parrett, K. G., Mehari, T., Jones, J. L., and Brand, J. J. (1988) *FEBS Lett. 228*, 268−272.
11. Parrett, K. G., Mehari, T., and Golbeck, J. H. (1990) *Biochim. Biophys. Acta 1015*, 341−352.
12. Li, N., Zhao, J. D., Warren, P. V., Warden, J. T., Bryant, D. A., and Golbeck, J. H. (1991) *Biochemistry 30*, 7863−7872.
13. Zhao, J., Snyder, W. B., Mühlenhoff, U., Rhiel, E., Warren, P. V., Golbeck, J. H., and Bryant, D. A. (1993) *Mol. Microbiol. 9*, 183−194.
14. Chitnis, P. R., Reilly, P. A., Miedel, M. C., and Nelson, N. (1989) *J. Biol. Chem. 264*, 18374−18380.
15. Bryant, D. A., Rhiel, E., de Lorimier, R., Zhou, J., Stirewalt, V. L., Gasparich, G. E., Dubbs, J. M., and Snyder, W. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., Ed.) pp 1−9, Kluwer Academic, Dordrecht, The Netherlands.
16. Yu, L., Golbeck, J. H., Zhao, J. D., Schluchter, W., Mühlenhoff, U., and Bryant, D. A. (1992) in *Current Research in Photosynthesis* (Murata, N., Ed.) pp 565−568, Kluwer Academic, Dordrecht, The Netherlands.
17. Rousseau, F., Sétif, P., and Lagoutte, B. (1993) *EMBO J. 12*, 1755−1765.
18. Mühlenhoff, U., Zhao, J., and Bryant, D. A. (1996) *Eur. J. Biochem. 235*, 324−331.
19. Rogers, L. J. (1987) in *The Cyanobacteria* (Fay, P., and van Baalen, C., Eds.) pp 35−61, Elsevier, Amsterdam, The Netherlands.
20. Straus, N. A. (1994) in *The Molecular Biology of Cyanobacteria* (Bryant, D. A., Ed.) pp 731−750, Kluwer Academic, Dordrecht, The Netherlands.
21. Barth, P., Lagoutte, B., and Sétif, P. (1998) *Biochemistry 37*, 16233−16241.
22. Falzone, C. J., Kao, Y.-H., Zhao, J., Bryant, D. A., and Lecomte, J. T. J. (1994) *Biochemistry 33*, 6052−6062.
23. Golbeck, J. H. (1994) in *The Molecular Biology of Cyanobacteria* (Bryant, D. A., Ed.) pp 319−360, Kluwer Academic, Dordrecht, The Netherlands.
24. Klukas, O., Schubert, W. D., Jordan, P., Krauss, N., Fromme, P., Witt, H. T., and Saenger, W. (1999) *J. Biol. Chem. 274*, 7351−7360.
25. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, Vol. 1, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Hanahan, D. (1983) *J. Mol. Biol. 166*, 557−580.
27. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
28. Laemmli, U. K. (1970) *Nature 227*, 680−685.
29. MacLaughlin, K. L. (1997) in *Chemistry Department*, p 122, The Pennsylvania State University, University Park, PA.
30. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
31. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
32. Schellman, J. A. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 115−137.
33. Dill, K. (1990) *Biochemistry 29*, 7133−7155.
34. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
35. Live, D. H., Davis, D. G., Agosta, W. C., and Cowburn, D. (1984) *J. Am. Chem. Soc. 106*, 1939−1941.
36. Falzone, C. J., Kao, Y.-H., Zhao, J., MacLaughlin, K. L., Bryant, D. A., and Lecomte, J. T. J. (1994) *Biochemistry 33*, 6043−6051.
37. Bodenhausen, G., and Ruben, D. G. (1980) *Chem. Phys. Lett. 69*, 185−189.
38. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
39. Neri, D., Otting, G., and Wüthrich, K. (1990) *J. Am. Chem. Soc. 112*, 3663−3665.
40. Billeter, M., Neri, D., Otting, G., Qian, Y. Q., and Wüthrich, K. (1992) *J. Biomol. NMR 2*, 257−274.
41. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., and Skelton, N. J. (1996) *Protein NMR Spectroscopy. Principles and Practice*, Academic Press, San Diego, CA.
42. Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) *J. Magn. Reson., Ser. B 101*, 333−337.
43. Muhandiram, D. R., and Kay, L. E. (1993) *J. Magn. Reson., Ser. B 102*, 317−321.
44. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
45. Bax, A., and Ikura, M. (1991) *J. Biomol. NMR 1*, 99−104.
46. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson. 99*, 201−207.
47. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
48. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
49. Brünger, A. T. (1992) *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
50. Nilges, M., Kuszewski, J., and Brünger, A. T. (1991) *Computational Aspects of the Study of Biological Macromolecules by NMR*, Plenum Press, New York.
51. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
52. Vriend, G. (1990) *J. Mol. Graph. 8*, 52−56.
53. Vriend, G., and Sander, C. (1993) *J. Appl. Crystallogr. 26*, 47−60.
54. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) *Biochemistry 33*, 2142−2150.
55. Chen, Y. J., Lin, S. C., Tzeng, S. R., Patel, H. V., Lyu, P. C., and Cheng, J. W. (1996) *Proteins: Struct., Funct., Genet. 26*, 465−471.
56. Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R. (1996) *J. Mol. Biol. 264*, 1132−1144.
57. Schmid, F. X. (1990) in *Protein Structure. A Practical Approach* (Creighton, T. E., Ed.) pp 251−285, IRL Press, New York.

58. Lim, W. A., Fox, R. O., and Richards, F. M. (1994) *Protein Sci. 3*, 1261−1266.
59. Viguera, A. R., Blanco, F. J., and Serrano, L. (1995) *J. Mol. Biol. 247*, 670−681.
60. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1995) *Biochemistry 34*, 868−878.
61. Grantcharova, V. P., and Baker, D. (1997) *Biochemistry 36*, 15685−15692.
62. Knapp, S., Mattson, P. T., Christova, P., Berndt, K. D., Karshikoff, A., Vihinen, M., Smith, C. I., and Ladenstein, R. (1998) *Proteins: Struct., Funct., Genet. 31*, 309−319.
63. Plaxco, K. W., Guijarro, J. I., Morton, C. J., Pitkeathly, M., Campbell, I. D., and Dobson, C. M. (1998) *Biochemistry 37*, 2529−2537.
64. Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D., and Dobson, C. M. (1998) *J. Mol. Biol. 276*, 657−667.
65. Robertson, A. D., and Murphy, K. P. (1997) *Chem. Rev. 97*, 1251−1267.
66. Shen, G., and Bryant, D. A. (1995) *Photosynth. Res. 44*, 51−53.
67. Vassiliev, I. R., Jung, Y. S., Yang, F., and Golbeck, J. H. (1998) *Biophys. J. 74*, 2029−2035.
68. Sonoike, K., Hatanaka, H., and Katoh, S. (1993) *Biochim. Biophys. Acta 1141*, 52−57.
69. Weber, N., and Strotmann, H. (1993) *Biochim. Biophys. Acta 1143*, 204−210.
70. Lagoutte, B., Barth, P., and Sétif, P. (1995) in *Photosynthesis: from Light to Biosphere* (Mathis, P., Ed.) pp 71−74, Kluwer Academic, Amsterdam, The Netherlands.
71. Sétif, P., Hanley, J., Barth, P., Bottin, H., and Lagoutte, B. (1995) in *Photosynthesis: From Light to Biosphere* (Mathis, P., Ed.) pp 23−28, Kluwer Academic, Amsterdam, The Netherlands.
72. Sun, J., Xu, Q., Chitnis, V. P., Jin, P., and Chitnis, P. R. (1997) *J. Biol. Chem. 272*, 21793−21802.
73. Feng, S., Chen, J., Yu, H., Simon, J., and Schreiber, S. (1994) *Science 266*, 1241−1247.
74. Xu, Q., Guikema, J. A., and Chitnis, P. R. (1994) *Plant Physiol. 106*, 617−624.
75. von Leoprechting, A., Horth, P., Haehnel, W., Schilz, E., and Mühlenhoff, U. (1998) *Anal. Biochem. 262*, 110−121.
76. Golbeck, J. H., and Bryant, D. A. (1991) in *Light Driven Reactions in Bioenergetics* (Lee, C. P., Ed.) pp 83−177, Academic Press, New York.
77. Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.
78. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.

BI9910373