N-Terminal Truncations of Manganese Stabilizing Protein Identify Two Amino Acid Sequences Required for Binding of the Eukaryotic Protein to Photosystem II and Reveal the Absence of One Binding-Related Sequence in Cyanobacteria†

Hana Popelkova,‡ Michael M. Im,‡ and Charles F. Yocum*,‡,§

Department of Molecular, Cellular, and Developmental Biology and Department of Chemistry, The University of Michigan, Ann Arbor, Michigan 48109-1048

Received March 22, 2002; Revised Manuscript Received May 17, 2002

ABSTRACT: Manganese stabilizing protein (MSP) is an intrinsically disordered extrinsic subunit of photosystem II that regulates the stability and kinetic performance of the tetranuclear manganese cluster that oxidizes water to oxygen. An earlier study showed that deletion of the ^{1}E−^{3}G domain of MSP caused no loss of activity reconstitution, whereas deletion of the ^{4}K−^{10}E domain reduced binding of the protein from 2 to 1 mol of MSP/mol of photosystem II and lowered activity reconstitution to about 50% of the control value [Popelkova et al. (2002) *Biochemistry 41*, 2702−2711]. In this work we present evidence that deletion of 13 or 14 amino acid residues from the MSP N-terminus (mutants ΔS13M and ΔK14M) does not interfere either with functional binding of one copy of MSP to photosystem II or with reconstitution of oxygen evolution activity to 50% of the control level. Both of these mutants exhibit nonspecific binding to photosystem II at higher protein concentrations. Truncation of the MSP sequence by 18 amino acids (mutant ΔE18M), however, causes a loss of protein binding and activity reconstitution. This result demonstrates that the N-terminal domain ^{15}T−^{18}E is required for binding of at least one copy of MSP to photosystem II. Analyses of CD spectra reveal changes in the structure of ΔE18M (loss of β-sheet, gain of unordered structure). Use of the information gained from these experiments in analyses of N-terminal sequences of MSP from a number of species indicates that higher plants and algae possess two recognition domains that are required for MSP binding to PSII, whereas cyanobacteria lack the first N-terminal domain found in eukaryotes. This may explain the absence of a second copy of MSP in the crystal structure of PSII from *Synechococcus elongatus* [Zouni et al. (2001) *Nature 409*, 739−743].

The water oxidizing activity of photosystem II (PSII)[1] takes place in the oxygen evolving complex (OEC) of cyanobacteria, algae, and higher plants and generates O_2, protons, and electrons. Substrate water binds to the redox-active site of the OEC, which cycles through five distinct redox states known as S-states (S_n, where $n = 0-4$) (*1*). Optimal functioning of the OEC depends on the presence of inorganic cofactors (4 Mn, 1 Ca^{2+}, and 1 Cl^{-}) that are ligated to the intrinsic protein domain of PSII (*2*). Three extrinsic proteins (17, 23, and 33 kDa) that are bound to the intrinsic domain of PSII regulate activity and stability of the inorganic cofactors (*3*). The 33 kDa extrinsic protein, also known as manganese stabilizing protein, or MSP, affects the properties of the tetranuclear manganese cluster. In the absence of MSP, high, nonphysiological, concentrations of Ca^{2+} and Cl^{-} are required for stable binding of Mn to PSII (*4*−*9*), as well as for sustained oxygen evolution activity, albeit at lower rates. An intimate role for MSP in Mn redox reactions is suggested by the finding that glutamate and/or aspartate residues of the protein are deprotonated during the S_1-S_2 transition (*10*).

The number of MSP copies bound per PSII reaction center is still a controversial question (*11*−*18*). Recent structural studies of PSII have employed electron microscopy for single particle analyses of the enzyme or X-ray diffraction applied to three-dimensional crystals. In the X-ray structure of a *Synechococcus elongatus* PSII preparation at 3.8 Å resolution, part of a single copy of MSP is shown as a region of β-sheet-rich structure protruding from the surface of the enzyme (*13*). Electron diffraction analyses of spinach preparations show a globular mass on the surface of PSII particles that is ascribed to a single copy of MSP (*11*), but there are alternate opinions as to the actual number of

† This research was supported by a grant to C.F.Y. from the National Science Foundation (MCB-0110455).

* Correspondence should be addressed to this author at the Department of Molecular, Cellular, and Developmental Biology. Telephone: (734) 647-0887. Fax: (734) 647-0884. E-mail: cyocum@umich.edu.

‡ Department of Molecular, Cellular, and Developmental Biology.

§ Department of Chemistry.

[1] Abbreviations: Bis-Tris, bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane; BSA, bovine serum albumin; CD, circular dichroism; Chl, chlorophyll; IPTG, isopropyl β-D-thiogalactopyranoside; MES, 2-(*N*-morpholino)ethanesulfonic acid; MSP, manganese stabilizing protein; OEC, O_2 evolving complex; PAGE, polyacrylamide gel electrophoresis; pET, plasmid for gene expression with T7 RNA polymerase; PS, photosystem; *psb*O, gene encoding precursor MSP; SDS, sodium dodecyl sulfate; sw-PSII, NaCl-washed photosystem II membranes depleted of 23 and 17 kDa extrinsic proteins; TOPO, plasmid for gene expression with T7 RNA polymerase; Tris, tris(hydroxymethyl)aminomethane; unrd, unordered structure; usw-PSII, urea salt-washed photosystem II membranes depleted of 33, 23, and 17 kDa extrinsic proteins; UV, ultraviolet; Δ, represents missing amino acid residues.

10.1021/bi020228u CCC: $22.00 © 2002 American Chemical Society
Published on Web 07/11/2002

subunits (one or two) that can occupy this space (*12*). Several biochemical studies to quantify bound MSP, or to determine the number of copies of the protein that are necessary to restore full activity, yield estimates of two copies of MSP bound per PSII reaction center (*14*−*18*).

The existing evidence suggests that the MSP−PSII interaction is stabilized by electrostatic interactions between positively and negatively charged amino acid residues (K, R, D, E) (*19*−*21*). Although it has been proposed that the intrinsic PSII subunits CP43 and cytochrome *b*-559 may function as binding sites for MSP (*22*), the large, extrinsic "E" loop of CP47 is the best characterized site of contact between PSII and MSP (*23*, *24*). Cross-linking by 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide (EDC) showed that the ^{1}E−^{76}K domain of MSP interacts with the ^{364}D−^{440}D domain of the CP47 E loop (*25*). Site-directed mutagenesis of amino acid residues in the E loop interferes with MSP binding to PSII, in agreement with the cross-linking result (*26*, *27*).

Manganese stabilizing protein exhibits an unusual behavior in solution that has led to its characterization as a natively unfolded, or intrinsically disordered, polypeptide (*28*). It is thermostable and, on the basis of far-UV CD analyses, is predicted to contain a relatively small amount of α-helix (<10%), <45% *β*-sheet, and about 50% random coil and unordered structure (*29*−*31*). Hydrodynamic data show that MSP can be modeled as a prolate ellipsoid, with an axial ratio of 4.2 (*32*). This is supported by recent small-angle X-ray scattering experiments, which yield an axial ratio of 4.8 (B. Barry, personal communication). When MSP is free in solution, Glu 246 and Lys 48 are cross-linked by EDC (*33*), which indicates that the N- and C-termini of MSP are in close proximity to one another, a result that might be expected for the shape predicted for MSP in solution.

Site-directed mutagenesis has been employed to obtain information about domains and residues of spinach MSP that are involved in binding and reconstitution of activity to the OEC. The C-terminal tripeptide, LEQ, has been shown to be important for quantitative assembly of MSP into PSII and restoration of high O_2 evolution activity (*34*); Leu 245 has been identified as the critical residue in this sequence (*31*). Also near the C-terminus, replacement of Val 235 with Ala yielded a temperature-sensitive binding defect in MSP, and V235A was shown to weaken binding of one of two copies of MSP to PSII (*16*, *17*). Seidler et al. (*3*, *35*) replaced the Asp residues at positions 109, 157, and 212. The mutants D109K, D157N, and D157K exhibited decreased binding affinities for PSII, whereas D109N and D212N retained normal binding characteristics. The MSP double mutant C28A,C51A, which lacks the single disulfide bridge present in this protein, exhibits normal binding and activity reconstitution (*36*). Wild-type levels of O_2 evolution activity were also observed with spinach mutants in which the Asp residue at position 9 in the spinach sequence was replaced by K (D9K) or N (D9N) (*37*). An alternative approach using truncation mutations at the N-terminus of MSP (*38*) has shown that removal of three amino acid residues (mutant ΔG3M) does not influence functional binding and activity but does induce nonspecific binding to PSII at higher protein concentrations. Truncation by ten N-terminal amino acids (mutant ΔE10M) reduced binding to PSII to 1 mol of ΔE10M/mol of PSII and lowered reconstituted activity to 50% of the control value, which indicates that the ^{4}K−^{10}E domain is required for binding of one copy of MSP to PSII. Both ΔG3M and ΔE10M exhibited greater overall protein compactness in comparison to wild-type MSP, although neither mutant exhibited a detectable change in secondary structure content (*38*).

In this paper, site-directed mutagenesis has been employed to produce three new truncation mutants of spinach MSP. The first two deletions, of 13 and 14 amino acid residues at the MSP N-terminus (ΔS13M and ΔK14M mutants), yielded proteins whose properties are similar to those reported for a less drastic truncation mutant, ΔE10M (*38*). In contrast to ΔE10M, however, both of these new mutants exhibited nonspecific binding to PSII at higher protein concentrations. A deletion mutation removing 18 N-terminal amino acids (ΔE18M mutant) was also engineered. This mutant behaved in a manner similar to that of the modified MSP created by removal of 18 residues by proteolysis with *Staphylococcus aureus* strain V8 protease, as described by Eaton-Rye and Murata (*39*). In agreement with those results, a characterization of ΔE18M shows that this mutant exhibits drastically reduced PSII binding and activity reconstitution. Combining this result with those obtained with other truncation mutants of MSP shows that the domain ^{15}T−^{18}E is necessary for binding of at least one copy of MSP to PSII. Use of these experimental results in analyses of MSP N-terminal amino acid sequence alignments from a number of eukaryotes and prokaryotes reveals that higher plants and algae possess two sequences that are necessary for MSP binding to PSII, whereas cyanobacteria lack one of the sequences found in eukaryotes.

MATERIALS AND METHODS

Mutations of psbO cDNA and Transformations of Escherichia coli Cells. The procedures for preparation of *psb*O cDNA and for transformation of *E. coli* cells were carried out as described in ref *38*, with minor modifications. For construction of cDNAs encoding the recombinant truncation mutants ΔS13M, ΔK14M, and ΔE18M, the sequences 5′-TCATATGAAGACATACCTCGAAGTC-3′, 5′-TCATATGACATACCTCGAAGTCAAAGGA-3′, and 5′-TCATATGGTCAAAGGAACTGGAACA-3′ were synthesized (Life Technology, Gibco BRL custom primers) on the basis of the sense strand of the MSP N-terminal coding region. As an antisense strand of the C-terminal coding region, the synthetic oligonucleotide 5′-TGAATTCCTTTTATTGCTCAAGTTGTGC-3′ was used. The N-terminal oligonucleotides contained an *Nde*I endonuclease restriction site, and the C-terminal oligonucleotide contained an *Eco*RI endonuclease restriction site three nucleotides after the stop codon for *psb*O. The mutant *psb*O cDNAs were subcloned into pCR T7/CT TOPO (Invitrogen) or pET 11-a (Novagen) translation vectors. The TOPO *psb*O constructs for ΔS13M and ΔK14M were inserted into BL21(DE3)pLysS *E. coli* cells, and the pET *psb*O construct for ΔE18M was inserted into BL21(DE3)pLysE *E. coli* cells. All mutations were confirmed by DNA sequencing. The integrity of each N-terminal recombinant mutant was also confirmed by amino acid sequencing (Edman degradation), and the expected N-terminal sequences MKTYL (ΔS13M), MTYLE (ΔK14M), and MVKGT (ΔE18M) were found (data not shown).

Overexpression and Purification of Recombinant MSP's. Recombinant wild-type MSP overexpressed in *E. coli* and purified from inclusion bodies is functionally indistinguishable from protein extracted from PSII (*40*). All truncation mutants were overexpressed in *E. coli*, and inclusion bodies were isolated as described in refs *38* and *40*, except that 0.5 mM IPTG was used for overexpression of N-terminal deletion mutants. Recombinant mutants were extracted from inclusion bodies and purified as described in ref *38*, with the modification that overnight incubation at 4 °C in 3 M urea, 5% betaine (w/v), 20 mM Tris (pH 8), and 5 mM NaCl was used for solubilization. The recombinant proteins were loaded onto a Pharmacia Resource Q column equilibrated with the same buffer, and a linear gradient (5−250 mM NaCl) was applied. Recombinant MSP eluted between 120 and 150 mM NaCl. One chromatographic step was sufficient to yield protein of high purity.

Reconstitution of PSII with MSP and Functional Analysis. Photosystem II membranes [250 Chl/reaction center (*17*, *36*)] were isolated, and extrinsic proteins were extracted as described in ref *38*. Reconstitution experiments and activity assays employed the procedures described in refs *36* and *38*, except that usw-PSII and MSP were mixed together in a solution containing 37 mM MES (pH 6), 10 mM $CaCl_2$, 80 mM NaCl, 0.1 mg of BSA/mL, 0.3 M sucrose, and 2% betaine (w/v), and 600 μM, rather than 300 μM, DCBQ was present in the assay buffer during O_2 rate measurements. The extent of rebinding of MSP to PSII was determined as described earlier (*36*, *38*). The MSP content of a sw-PSII sample was taken as the control (100% MSP bound), except for the experiments using intact PSII, where the level of natively bound MSP was taken as the control. Coomassie staining of MSP is linear with protein concentration (data not shown). Some decrease in dye binding is likely with the truncation mutants, but the most substantial of these (ΔE18M) would produce an error of about 7% in the estimation of MSP concentration on Coomassie-stained gels.

Structural Properties of MSP Mutants. Spectroscopic characterizations of MSP samples were carried out as described in ref *38*; conditions are given in the appropriate figure legends. The CD spectrometer was calibrated with (+)-10-camphorsulfonic acid (1 mg/mL in H_2O) (*41*). Estimations of protein secondary structure content from far-UV CD spectra were obtained using CONTIN/LL and CDSSTR methods in the CDPro package downloaded from the web page http://lamar.colostate.edu/~sreeram/CDPro. For detailed information about methods and basis sets, see ref *42*. A Superose-12 column on a Pharmacia FPLC system was used for size exclusion chromatography to estimate molecular masses of the truncated forms of MSP (*31*, *38*).

Amino Acid Sequence Alignment. The complete amino acid sequence of spinach MSP was located using ENTREZ (http://www.ncbi.nlm.nih.gov/) and entered into PSI- and PHI-BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) to identify the other known MSP sequences. The sequences of various organisms, presented in Table 2, were entered into the input field of BCM Search Launcher, and the Clustal W 1.8 program was used for sequence alignment (http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html).

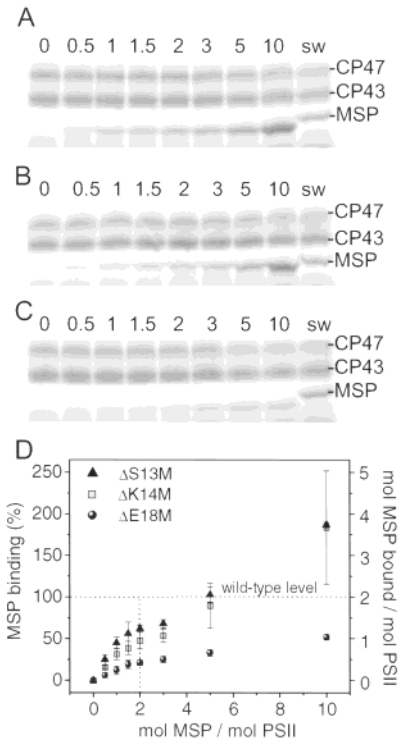

FIGURE 1: Binding of ΔS13M, ΔK14M, and ΔE18M to PSII. The Coomassie blue stained SDS−PAGE gel shows usw-PSII samples reconstituted with ΔS13M (A), ΔK14M (B), and ΔE18M (C), respectively, for 1 h at room temperature. The moles of MSP per mole of PSII used for reconstitution is indicated at the top of the lanes; sw, control sw-PSII. (D) Binding curves for usw-PSII reconstituted with ΔS13M, ΔK14M, and ΔE18M, respectively. SDS−PAGE gels of reconstituted samples were analyzed by densitometry. Binding efficiency of ΔS13M, ΔK14M, and ΔE18M was expressed as a function of added MSP (mol/mol of PSII; *x* axis), both as a fraction of MSP bound to the sw-PSII control (100%, left *y* axis) and as the amount of bound moles of MSP per mole of PSII (right *y* axis). Each point is the average of three separate experiments, and vertical bars represent the standard deviation.

RESULTS

In the N-terminal sequence of spinach MSP (EGGKRL-TYDEIQSKTYLE-), an earlier report identified a domain, $^{4}K-^{10}E$, as essential for binding of one copy of MSP to PSII (*38*). To attempt to locate the amino acid sequence needed for binding of the second copy of MSP to PSII, new mutants were prepared by substituting a translation initiation codon (Met residue) at positions 13, 14, and 18 of the spinach protein. The N-terminal sequences of these mutant proteins are given in Materials and Methods.

Functional Characterization of ΔS13M, ΔK14M, and ΔE18M. The mutation ΔS13M eliminated the first 13 amino acid residues from MSP's N-terminus. Figure 1A presents results of an SDS−PAGE experiment to characterize rebinding of ΔS13M to usw-PSII membranes. Several points are revealed by these data. First, it is clear that, by comparison with native MSP bound to sw-PSII membranes, ΔS13M exhibits a faster migration on the gel. This is probably due to a reduction in size of the protein. Results from a quantitative evaluation of the gel are shown in Figure 1D. As can be seen, ΔS13M binding differs from that of wild-type MSP (indicated by the dotted lines). While wild-type binding saturates at 2 mol of MSP/mol of PSII, equal to the

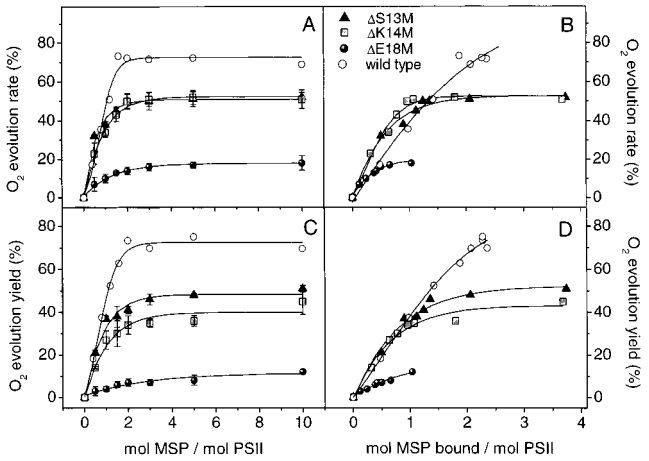

FIGURE 2: Recovery of O_2 evolution activity by usw-PSII membranes reconstituted with ΔS13M, ΔK14M, and ΔE18M, respectively. (A) O_2 rate and (C) O_2 yield assayed during 1 and 4 min of continuous illumination, respectively, plotted as a function of moles of MSP added to reconstitution mixtures. (B) O_2 rate and (D) O_2 yield as a function of MSP bound to PSII. Samples were assayed after reconstitution incubation for 1 h at room temperature. The residual activity of usw-PSII was subtracted from the activities of reconstituted samples. 100% corresponds to the activity of control sw-PSII. Points are averages of three experiments, and the vertical bars at each point (panels A and C) give the standard deviation. For wild type, the data are redrawn from ref *38*.

MSP content of the control sw-PSII sample, ΔS13M binding increased gradually and did not attain saturation at the highest protein concentrations added to reconstitution mixtures. About 1.2 mol of ΔS13M/mol of PSII was bound when 2 mol of ΔS13M/mol of PSII was used for reconstitution, and about 4 mol of ΔS13M/mol of PSII was bound when 10 mol of ΔS13M/mol of PSII was present in reconstitution mixtures (Figure 1D). Recovery of O_2 evolution activity (rate and yield assays) by reconstitution with ΔS13M attained about 50% of the control value at saturation (Figure 2). This ΔS13M activity represents about 70% of the reconstitution activity of the wild-type protein (saturation at 70% of the control value upon binding of 2 mol of MSP/mol of PSII; see Figure 2). Panels A and C of Figure 2 indicate that about 2 mol of ΔS13M/mol of PSII in reconstitution mixtures was required to attain saturation of functional protein binding, as monitored by recovery of activity. The relationship between actual binding of MSP and reconstitution of activity is shown in Figure 2B,D, where oxygen evolution activity is plotted as a function of ΔS13M bound to PSII. As these figures show, binding of 1 mol of ΔS13M/mol of PSII produced maximal recovery of activity, despite the ability of this mutant to bind nonspecifically (up to 4 mol of protein/mol of PSII; compare Figures 1 and 2).

The role of the positively charged lysine residue at position 14 in the spinach MSP sequence was characterized using the ΔK14M mutant. The effects of this mutation on MSP rebinding and activity reconstitution are shown in Figures 1 and 2. On the SDS gel (Figure 1B), ΔK14M migrates slightly faster than wild type, and rebinding assays show that the behavior of ΔK14M is similar to that of ΔS13M (Figure 1D). Like ΔS13M, ΔK14M binding fails to attain saturation. About 1 mol of ΔK14M was bound to PSII upon addition of 2 mol of ΔK14M/mol of PSII to the reconstitution mixture, and almost 4 mol of ΔK14M/mol of PSII was bound when 10 mol of ΔK14M/mol of PSII was present in the reconstitution mixture. The activity assays with ΔK14M yielded results very similar to those obtained with ΔS13M (Figure 2), the only difference being a 10% lower O_2 evolution yield (Figure 2C,D). On the basis of these results, the positively charged lysine residue, ^{14}K, on the N-terminus of spinach MSP is not required for functional binding of one copy of MSP to PSII.

Deletion of up to 14 amino acid residues from the spinach MSP N-terminus fails to affect the ability of the truncation mutants to bind one copy of MSP to PSII and to partially restore the function of the OEC. On the basis of these data, a more drastic truncation of MSP was engineered, to produce the mutant ΔE18M, which imitates the effect of protease truncation of MSP reported by Eaton-Rye and Murata (*39*). Like the other truncation mutants of MSP, ΔE18M migrated faster on the SDS−PAGE gel (Figure 1C), but the PSII binding affinity of this protein was significantly impaired, compared to the affinities of ΔS13M and ΔK14M (Figure 1D). About 0.4 mol of ΔE18M was bound when 2 mol of ΔE18M/mol of PSII was added to reconstitution mixtures, and only 1 mol of ΔE18M/mol of PSII was bound at the highest protein concentration used for reconstitution (10 mol of ΔE18M/mol of PSII) (Figure 1D). Oxygen evolution rate and yield assays also revealed a noticeable decrease in the functional capacity of this mutant (Figure 2). Maximum binding of ΔE18M (about 1 mol/mol of PSII) reconstituted about 20% of the control O_2 evolution rate but only 10% of the O_2 yield (Figure 2). These data show that elimination of the first 18 amino acid residues from the N-terminus of spinach MSP results in a severe defect in binding and activity reconstitution by ΔE18M.

The ability of all three truncation mutants to bind nonspecifically to PSII was tested in experiments using intact PSII membranes in which MSP binding sites are occupied by the native protein. The results of these experiments show that ΔS13M and ΔK14M will rebind nonspecifically to an intact preparation (about 1 mol of MSP/mol of PSII binds upon incubation of intact PSII with 10 mol of mutant/mol of PSII), while ΔE18M exhibits little, if any, nonspecific binding (data not shown).

Structural Properties of ΔS13M, ΔK14M, and ΔE18M in Solution. As a natively unfolded or intrinsically disordered protein (*28*), MSP exhibits a higher apparent molecular mass (37−41 kDa) on size exclusion chromatography than would be predicted on the basis of its molecular mass (26.5 kDa) calculated from the amino acid sequence derived from the DNA sequence (*28*, *34*, *38*, *43*). In comparison to wild-type MSP (37.2 kDa), all deletion mutants exhibited significantly lower molecular masses on a calibrated gel filtration column; 32.7 kDa (ΔS13M) and 32.8 kDa (ΔK14M and ΔE18M). These diminished masses, relative to the wild-type protein, are still greater (by about 8 kDa) than the masses calculated from the appropriate amino acid sequences (data not shown).

Manganese stabilizing protein has a number of aromatic amino acid residues in its sequence; in spinach, a single Trp residue is located at position 241 in the C-terminus (*31*), and 8 Tyr and 13 Phe residues are distributed throughout the MSP sequence. Any change in the environment of these residues mirrors structural changes in the C-terminus (Trp) or in the overall structure of MSP (Tyr, Phe). Changes in the environment of aromatic amino acid residues can be reflected in the amplitudes of characteristic features of UV

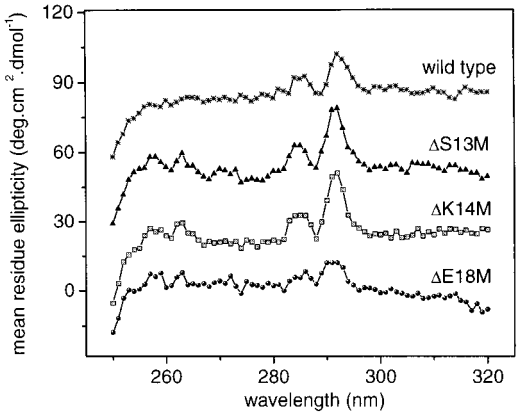

FIGURE 3: Near-UV CD spectra (protein solution minus buffer) of wild type, ΔS13M, ΔK14M, and ΔE18M. Each spectrum is the average of 20 scans. Experimental conditions were as follows: time constant, 1 s; bandwidth, 1 nm; scan step, 1 nm; room temperature; scan width, 320−250 nm; path length, 1 cm; sample volume, 1 mL.

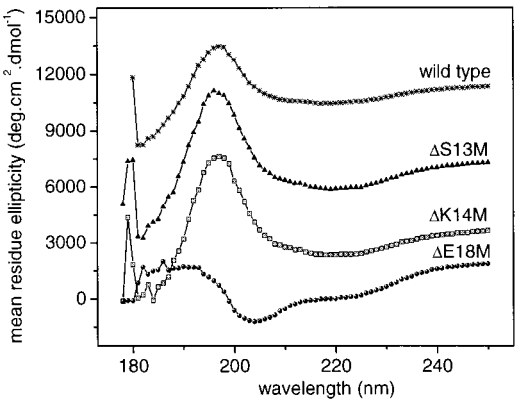

FIGURE 4: Far-UV CD spectra (protein solution minus buffer) of wild type, ΔS13M, ΔK14M, and ΔE18M. Each spectrum is the average of 20 scans. Experimental conditions were as follows: time constant, 1 s; bandwidth, 1 nm; scan step, 1 nm; room temperature; scan width, 250−178 nm; path length, 1 mm; sample volume, 300 μL.

absorption and near-UV CD spectra. In general, an increase in the hydrophobicity of the environment surrounding aromatic residues will produce an increase in the intensity of these spectral features, and imposition of an increased hydrophilic environment of these residues is indicated by the disappearance of these same spectral features. The UV absorption spectrum of wild-type MSP has a characteristic shoulder at 293 nm, arising from tryptophan (*44*), so the UV spectrum of the protein can be a useful tool for monitoring structural changes at the C-terminus of MSP. The UV spectra of the deletion mutants were examined. For ΔS13M and ΔK14M, the characteristic shoulder at 293 nm was slightly more distinguishable than that of wild-type MSP, because it was shifted slightly to the red. In the case of ΔE18M, the elimination of additional amino acids from the MSP N-terminus caused a decline in Trp absorption and, thus, a less prominent 293 nm shoulder as compared to the wild-type protein (data not shown).

A change in the overall tertiary structure of MSP, which affects the environment of all aromatic amino acid residues, was monitored by near-UV CD spectroscopy (*45*). The near-UV CD spectrum of wild-type MSP has characteristic peaks at 285 and 292 nm (Figure 3) assigned to tyrosine and tryptophan, respectively (*45*). The results obtained from these deletion mutants (Figure 3) are consistent with the results from UV absorption measurements; ΔS13M and ΔK14M exhibited well-defined Tyr and Trp peaks at 285 and 292 nm, respectively, which are more prominent than in the wild-type protein. Small peaks at 258 and 263 nm, probably from phenylalanine (*45*), could also be distinguished. All of these peaks were diminished in the near-UV CD spectrum of ΔE18M to levels below those detected in the wild-type protein (Figure 3).

To monitor protein secondary structure, far-UV CD spectra of wild-type MSP and the deletion mutants were obtained (Figure 4). Spectra of ΔS13M and ΔK14M exhibit higher amplitudes of mean residue ellipticity, but the same shapes, as the spectrum of the wild-type protein. The spectrum of ΔE18M, however, showed a somewhat different shape than that of wild-type protein or of the deletion mutants. These spectra were analyzed with the CONTIN/LL and CDSSTR programs (see Materials and Methods). Basis sets 1, 3, and 4 and basis sets 6 and 7 were used for MSP secondary structure prediction from CD spectra. The latter basis sets are probably more accurate for secondary structure estimation of MSP because they contain CD spectra of five unfolded proteins (*42*). Analyses of far-UV CD spectra did not reveal any significant differences in the estimated secondary structure contents of wild type, ΔS13M, and ΔK14M (Table 1), similar to the results obtained with the ΔG3M and ΔE10M deletion mutants (see ref *38*). Analysis of the spectrum of ΔE18M, however, showed that the apparent changes in the amplitudes of mean residue ellipticity (Figure 4) reflected a decrease in the estimated content of β-sheet structure, and an increase in the estimated content of turns and unordered secondary structure, relative to the secondary structure content of the wild-type protein (Table 1).

Table 1: Secondary Structure Predictions for Wild Type and Deletion Mutants of MSP Based on Analysis of Far-UV CD Spectra

	basis sets 1, 3, 4, 6, and 7[a]				basis sets 6 and 7 only			
protein	α-helix	β-sheet	turn + unrd	total	α-helix	β-sheet	turn + unrd	total
wild type	4	40	55	99	3	41	56	100
ΔS13M	5	40	54	99	3	39	57	99
ΔK14M	5	42	52	99	3	41	56	100
ΔE18M	6	35	58	99	4	34	61	99

[a] For a detailed basis set description, see ref *42*; numbers are averages of results obtained from the basis sets shown and from both CONTIN/LL and CDSSTR methods.

DISCUSSION

More than 50% of the amino acid residues in the N-terminus of MSP are either conserved or conservatively substituted across the 15 eukaryotic species for which amino acid sequence data are available (see Table 2). Truncation of the first three N-terminal amino acid residues of spinach MSP (mutant ΔG3M) has no effect on MSP function, although an increase in nonspecific binding of ΔG3M to PSII at higher protein concentrations was detected (*38*). In contrast, deletion of the first 10 amino acid residues (mutant ΔE10M) causes loss of binding of one copy of MSP to PSII (*38*). Characterization of the truncation mutants ΔS13M,

Table 2: Alignment of the N-Terminal Sequence of Mature MSP[a]

Species		N-terminal sequence
Eukaryotes -	Spinach	E G G - K R L T Y D E I Q S K T Y L E V K
-Higher plants	Tobacco	E G V P K R L T F D E I Q S K T Y M E V K
	Tomato	E G V P K R L T Y D E I Q S K T Y M E V K
	Pea	E G A P K R L T F D E I Q S K T Y L E V K
	Rice	E G V P R R L T F D E I Q S K T Y M E V K
	Wheat	E G A P K R L T F D E I Q S K T Y M E V K
	Potato	E G V P K R L T F D E I Q S K T Y M E V K
	Arachis	E G A P K R L T F D E I Q S K T Y L E V K
	Bruguiera	E G V P K R L T Y D E I Q S K T Y L E V K
	Fritillaria	E G V P K R L T F D E I Q S K T Y M E V K
	Brassica	E G A P K R L T Y D E I Q S K T Y M E V K
	Arabidopsis	E G A P K R L T Y D E I Q S K T Y M E V K
-Unicellular	Euglena	- - - - A S L T Y D E L Q S L S Y L E V K
organisms	Volvox	- - - - - - - T Y D E L Q G L T Y L Q V K
	Chlamydomonas	- - - - - - - T F D E I Q G L T Y L Q V K
Prokaryotes	Cyanothece	- - - - - - - - - V N P Q D L T Y D E I L
	Synechocystis	- - - - - - - - - V D K S Q L T Y D D I V
	Synechococcus PCC7942	- - - - - - - - - A D L G T L T Y D Q I K
	Synechococcus elongatus	- - - - - - - - - - A K Q T L T Y D D I V
	Anabaena	- - - - - - - - - - - R D I L T Y E Q I R

[a] Missing residues are indicated by a dash; boxes mark highly conserved sequence motifs at positions 7−10 and positions 15 and 16 in the amino acid sequence of spinach MSP.

ΔK14M, and ΔE18M allows us to identify the second domain of the MSP N-terminus that is required for functional binding of at least one of two copies of this protein to PSII.

Extending MSP N-terminal truncations from ^{10}E to 13 or 14 amino acid residues (ΔS13M and ΔK14M mutants) produces no change in the functional binding of these proteins to PSII (Figure 1). For both mutants, addition of about 2 mol of MSP/mol of PSII to reconstitution mixtures was necessary to attain functional binding of 1 mol of MSP/mol of PSII, indicating that the binding affinities of both mutants were lower than that of wild-type MSP. Protein binding curves did not saturate, however, and about 4 mol of MSP/mol of PSII was bound when 10 mol of either MSP mutant/mol of PSII was included in reconstitution mixtures. Possible artifacts from ineffective washing of PSII membranes after reconstitution, or from protein aggregation prior to reconstitution, were tested by more extensive washing of reconstituted PSII membranes and by centrifugation of MSP solutions before reconstitution, respectively. These treatments failed to eliminate the nonspecific-binding phenomenon, which was previously observed with ΔG3M (*38*). Because activity assays employing ΔS13M and ΔK14M showed that binding of 1 mol of MSP/mol of PSII was sufficient to saturate restoration of oxygen evolution activity at about 50% of the control value (Figure 2B,D), nonspecific binding by ΔS13M or ΔK14M (up to 3 mol of MSP/mol of PSII) occurs only after functional binding sites are occupied by either mutant. The nonspecific binding sites are unrelated to functional MSP binding sites; binding experiments with intact PSII membranes show that both ΔS13M and ΔK14M bind to these preparations, which already possess specifically bound native MSP. Although we cannot explain the cause of nonspecific binding by the mutants, it is not unique to our experiments. Motoki et al. (*46*) also showed, for a number of *S. elongatus* MSP mutants, a similar nonspecific-binding phenomenon. On the basis of the data presented in Figures 1 and 2, we conclude that both ΔS13M and ΔK14M are defective in specific binding of one of two copies of MSP to PSII and that these mutants, like ΔG3M, exhibit nonspecific PSII binding. The behavior of these mutants also further defines the MSP domain ^{4}K−^{10}E as an essential determinant that is required for specific binding of one copy of MSP to PSII (*38*).

A substantial modification of MSP binding to PSII was observed upon truncation of the first 18 amino acid residues of the protein's N-terminus (mutant ΔE18M). This truncation produces a protein that should, in theory, be very similar to the protease-modified MSP described by Eaton-Rye and Murata (*39*), in which removal of the 18 N-terminal amino acids abolished MSP binding to PSII. Our data show that ΔE18M exhibits significantly lower binding affinity to PSII [0.4 mol of ΔE18M/mol of PSII was bound when 2 mol of protein/mol of PSII was included in the reconstitution mixture; binding of 1 mol of ΔE18M/mol of PSII occurred upon inclusion of 10 mol of protein/mol of PSII in the reconstitution mixture (see Figure 1D)]. With about 1 mol of ΔE18M/mol of PSII bound to PSII, O_2 evolution rates and yields saturated at 18% and 11% of the control values, respectively (Figure 2B,D), and little nonspecific binding of ΔE18M to PSII was detected. These results, in reasonable agreement with those of Eaton-Rye and Murata (*39*), show conclusively that sequences within the first 18 amino acids

of the spinach MSP N-terminus are essential for functional PSII binding of two copies of MSP.

While secondary structure contents estimated for ΔS13M and ΔK14M were similar to the secondary structure of wild type (Table 1), analysis of far-UV CD spectra showed that ΔE18M gained some turns and unordered structural components at the expense of β-sheets (Table 1). The predicted loss of β-sheet structure in ΔE18M is consistent with the location of predicted secondary structure elements in the MSP sequence. The hypothetical structural model of MSP proposed by Bricker and Frankel (*9*) predicts β-sheet structures in the domains ^{4}K−^{9}D and ^{12}K−^{20}K, and an analysis of CD data, obtained from spinach MSP, yielded predictions of β-sheets in the sequences ^{5}R−^{8}Y and ^{15}T−^{19}V (*29*). The sequence ^{15}T−^{19}V was proposed to be the location of β-sheet in a consensus sequence of *psb*O derived from a multiple sequence alignment (*47*). All of these proposed β-sheet domains have been deleted in ΔE18M.

Characterization of the structural consequences of N-terminal deletions also revealed modifications to the solution tertiary structures of ΔS13M, ΔK14M, and ΔE18M. UV absorption (not shown) and near-UV CD spectra (Figure 3) indicate that aromatic amino acid residues of ΔS13M and ΔK14M are more solvent shielded than in the wild-type protein, whereas in ΔE18M these same residues are more solvent exposed (see Figure 3). Tyrosine is globally distributed in MSP, so the solution structures of ΔS13M and ΔK14M appear to be somewhat more compact than that of wild type. For ΔE18M, some loss of Tyr features in the near-UV CD spectrum would suggest that there has been a small reduction in overall hydrophobicity. For all three mutants, the hydrophobicity of the lone Trp near the C-terminus has been slightly enhanced (ΔS13M, ΔK14M) or reduced (ΔE18M) as compared to wild type. In solution, the C- and N-termini of MSP are believed to be in close proximity to one another (*33*), so it is possible that truncations at ^{13}S and ^{14}K, or at ^{18}E, affect this interaction and perturb the environment of Trp at position 241.

The data presented here and in ref *38* define two sequence domains of the MSP N-terminus that are necessary for its high-affinity binding to PSII. Elimination of the first domain, ^{4}KRLTYD^{10}E, reduces the number of copies of MSP that rebind to PSII from 2 to 1 (*38*). As we show in this paper, deletion of the second sequence, ^{15}TYL^{18}E, dramatically impairs binding of the other copy of MSP and reconstitution of activity. Although these data reveal the presence of two sequences that are required for binding of two copies of spinach MSP to PSII, a recent structural study arrived at a different conclusion regarding bound spinach MSP. Combining a theoretical model of MSP tertiary structure, on the basis of threading routines, with cryoelectron microscopy of a spinach PSII preparation has yielded a model of PSII structure in which a single copy of MSP is bound to eukaryotic PSII, in agreement with the single copy of MSP found in the 3.8 Å *S. elongatus* crystal structure (*48*). Our results provide an alternate explanation for the apparent discrepancy in MSP/PSII stoichiometries in eukaryotes and prokaryotes, i.e., why two copies are found in eukaryotes (*14*−*18*; this study). Table 2 presents all of the available MSP N-terminal sequences. The aligned sequences reveal a unique feature of cyanobacterial MSP's, namely, a truncation of the N-terminus that places the first binding-related sequence (TYDE) in alignment with the second binding-related sequence of eukaryotic MSP's. The alignments in Table 2 also show that, in unicellular eukaryotes, the N-terminus is truncated relative to higher plants but that these organisms retain the sequence TYDE of the first binding-related domain.

Only one study of MSP stoichiometries required for reconstitution in cyanobacteria is currently available (*46*), and in that investigation it is not entirely clear that binding of one copy of recombinant protein is sufficient to saturate PSII binding sites. From the results of Table 2, it is nevertheless apparent that cyanobacteria lack one of the sequences that is essential for binding of two copies of eukaryotic MSP to PSII. This conclusion is also apparent in the sequence alignments in ref *49*. An alternate alignment of these sequences is possible (*3*, *46*):

Spinach	E G G P K R L T Y D E I Q S K T Y L E V K G …
S. elongatus	A K Q T L T Y D D I - - - - - - - - - V G …

This shows that the cyanobacterial sequence (TYDD) aligns better with the first, rather than second, sequence of the eukaryotic protein. However, this alignment requires insertion of a gap (dashes) in the cyanobacterial sequence. This agrees with other published strategies for aligning MSP sequences from eukaryotes and cyanobacteria, which require introductions of gaps in sequences from one set of species or the other (*3*, *46*, *49*). This gap is probably a consequence of the absence of one of the binding recognition sequences in the cyanobacterial protein.

The possible absence of a second copy of MSP in cyanobacteria may signal a unique difference between the OEC of these organisms and that of eukaryotes. It has been suggested that cyanobacteria may possess another PSII subunit, along with MSP, that regulates the stability and activity of the OEC (*50*, *51*). If, in cyanobacteria, one copy of MSP has been replaced by another extrinsic subunit [for example, *psb*V (*50*, *51*)], this may account for the differences in effects of MSP deletion by mutagenesis that are observed in these organisms and in a unicellular eukaryote. In *Chlamydomonas*, MSP deletion produces nonphotoautotrophic cells in which accumulation of other PSII subunits is suppressed (*52*). The same deletion in *Synechocystis* 6803 produces cells that can survive by photoautotrophic growth, even though their capacity for oxygen evolution has been seriously impaired (*53*). Survival of the cyanobacterial Δ*psb*O strain would be consistent with proposals that another PSII subunit in these organisms also plays a role in stability and/or function of the OEC and is consistent as well with our data showing that deletion of one bound MSP from eukaryotic PSII does not cause a total loss of OEC function.

In conclusion, results of this work provide the identities of MSP N-terminal sequences (^{4}KRLT(Y/F)D^{10}E and ^{15}TY(L/M)^{18}E) that are required for binding of two copies of eukaryotic MSP to PSII. We cannot determine whether any of the charged residues in these sequences form salt bridges to the intrinsic subunits of PSII; a conservative interpretation of our results would be that these essential amino acid sequences target MSP to the appropriate docking sites on PSII. On the basis of alignments of MSP N-terminal sequences from higher plants, unicellular eukaryotes, and cyanobacteria, we conclude that one of these sequences is

missing from the latter organisms, and we propose that this may explain the presence of one, rather than two, copies of MSP in prokaryotes. N-Terminal deletions from spinach MSP yield proteins with somewhat reduced sizes (gel filtration), slightly enhanced or reduced hydrophobic environments around aromatic amino acid residues, and unmodified or slightly modified secondary structures. From these observations, we conclude that the MSP N-terminus may be necessary to maintain a less compact tertiary structure in solution. This contrasts with results from C-terminal modifications of MSP, where hydrodynamic and spectroscopic data reveal large changes in tertiary and secondary structure accompanied by greatly expanded overall size (*31*). Finally, our data support an earlier proposal that the N- and C-termini of MSP are in close contact (*33*). Truncations of MSP at the N-terminus affect the molar ellipticity of ^{241}W, near the C-terminus of MSP, consistent with a pronounced change in the solvent exposure of the C-terminal segment of the protein.

REFERENCES

1. Kok, B., Forbush, B., and McGloin, M. (1970) *Photochem. Photobiol. 11*, 457−475.
2. Debus, R. J. (1992) *Biochim. Biophys. Acta 1102*, 269−352.
3. Seidler, A. (1996) *Biochim. Biophys. Acta 1277*, 35−60.
4. Bricker, T. M. (1992) *Biochemistry 31*, 4623−4628.
5. Miyao, M., and Murata, N. (1984) *FEBS Lett. 170*, 350−354.
6. Ono, T. A., and Inoue, Y. (1983) *FEBS Lett. 164*, 255−259.
7. Miyao, M., Murata, N., Lavorel, J., Maison-Peteri, B., Boussac, A., and Etienne, A. L. (1987) *Biochim. Biophys. Acta 890*, 151−159.
8. Ono, T. A., and Inoue, Y. (1986) *Biochim. Biophys. Acta 850*, 380−389.
9. Bricker, T. M., and Frankel, L. K. (1998) *Photosynth. Res. 56*, 157−173.
10. Hutchison, R. S., Steenhuis, J. J., Yocum, C. F., Razeghifard, M. R., and Barry, B. A. (1999) *J. Biol. Chem. 274*, 31987−31995.
11. Nield, J., Orlova, E. V., Morris, E. P., Gowen, B., van Heel, M., and Barber, J. (2000) *Nat. Struct. Biol. 7*, 44−47.
12. Boekema, E., van Breemen, J. F. L., van Roon, H., and Dekker, J. P. (2000) *Biochemistry 39*, 2907−12915.
13. Zouni, A., Witt, H. T., Kern, J., Fromme, P., Krauss, N., Saenger, W., and Orth, P. (2001) *Nature 409*, 739−743.
14. Leuschner, C., and Bricker, T. M. (1996) *Biochemistry 35*, 4551−4557.
15. Xu, Q., and Bricker, T. M. (1992) *J. Biol. Chem. 267*, 25816−25821.
16. Betts, S. D., Ross, J. R., Pichersky, E., and Yocum, C. F. (1997) *Biochemistry 36*, 4047−4053.
17. Betts, S. D., Ross, J. R., Pichersky, E., and Yocum, C. F. (1996) *Biochemistry 35*, 6302−6307.
18. Seidler, A. (1994) *Protein Eng. 7*, 1277−1280.
19. Frankel, L. K., Cruz, J. A., and Bricker, T. M. (1999) *Biochemistry 38*, 14271−14278.
20. Miura, T., Shen, J. R., Takahashi, S., Kamo, M., Nakamura, E., Ohta, H., Kamei, A., Inoue, Y., Domae, N., Takio, K., Nakazoto, K., Inoue, Y., and Enami, I. (1997) *J. Biol. Chem. 272*, 3788−3798.
21. Seidler, A. (1996) *Eur. J. Biochem. 242*, 485−490.
22. Enami, I., Tohri, A., Kamo, M., Ohta, H., and Shen, J. R. (1997) *Biochim. Biophys. Acta 1320*, 17−26.
23. Putnam-Evans, C., and Bricker, T. M. (1994) *Biochemistry 33*, 10770−10776.
24. Enami, I., Satoh, K., and Katoh, S. (1987) *FEBS Lett. 226*, 161−165.
25. Odom, W. R., and Bricker, T. M. (1992) *Biochemistry 31*, 5616−5620.
26. Putnam-Evans, C., and Bricker, T. M. (1992) *Biochemistry 31*, 11482−11488.
27. Putnam-Evans, C., Burnap, R., Wu, J., Whitmarsh, J., and Bricker, T. M. (1996) *Biochemistry 35*, 4046−4053.
28. Lydakis-Simantiris, N., Hutchison, R. S., Betts, S. D., Barry, B. A., and Yocum, C. F. (1999) *Biochemistry 38*, 404−414.
29. Xu, Q., Nelson, J., and Bricker, T. M. (1994) *Biochim. Biophys. Acta 1188*, 427−431.
30. Shutova, T., Irrgang, K. D., Shubin, V., Klimov, V. V., and Renger, G. (1997) *Biochemistry 36*, 6350−6358.
31. Lydakis-Simantiris, N., Betts, S. D., and Yocum, C. F. (1999) *Biochemistry 38*, 15528−15535.
32. Zubrzycki, I. Z., Frankel, L. K., Russo, P. S., and Bricker, T. M. (1998) *Biochemistry 37*, 13553−13558.
33. Enami, I., Kamo, M., Ohta, H., Takahashi, S., Miura, T., Kusayanagi, M., Tanabe, S., Kamei, A., Motoki, A., Hirano, M., Tomo, T., and Satoh, K. (1998) *J. Biol. Chem. 273*, 4629−4634.
34. Betts, S. D., Lydakis-Simantiris, N., Ross, J. R., and Yocum, C. F. (1998) *Biochemistry 37*, 14230−14236.
35. Seidler, A., Röll, K., and Michel, H. (1992) in *Research in Photosynthesis* (Murata, N., Ed.) pp 409−412, Kluwer Academic Press, Dordrecht, The Netherlands.
36. Betts, S. D., Ross, J. R., Hall, K. U., Pichersky, E., and Yocum, C. F. (1996) *Biochim. Biophys. Acta 1274*, 135−142.
37. Seidler, A., Rutherford, A. W., and Michel, H. (1996) *Plant Mol. Biol. 31*, 183−188.
38. Popelkova, H., Im, M. M., D'Auria, J., Betts, S. D., Lydakis-Symantiris, N., and Yocum, C. F. (2002) *Biochemistry 41*, 2702−2711.
39. Eaton-Rye, J. J., and Murata, N. (1989) *Biochim. Biophys. Acta 977*, 219−226.
40. Betts, S. D., Hachigian, T. M., Pichersky, E., and Yocum, C. F. (1994) *Plant Mol. Biol. 26*, 117−130.
41. Johnson, W. C. (1990) *Proteins: Struct., Funct., Genet. 7*, 205−214.
42. Sreerama, N., and Woody, R. W. (2000) *Anal. Biochem. 287*, 252−260.
43. Tyagi, A., Hermans, J., Steppuhn, J., Jansson, Ch., Vater, F., and Herrmann, R. G. (1987) *Mol. Gen. Genet. 207*, 288−293.
44. Schmid, F. X. (1997) in *Protein Structure: A Practical Approach* (Creighton, T. E., Ed.) 2nd ed., pp 261−267, Oxford University Press, New York.
45. Kelly, S. M., and Price, N. C. (1997) *Biochim. Biophys. Acta 1338*, 161−185.
46. Motoki, A., Usui, M., Shimazu, T., Hirano, M., and Katoh, S. (2002) *J. Biol. Chem. 277*, 14747−14756.
47. De Las Rivas, J., and Heredia, P. (1999) *Photosynth. Res. 61*, 11−21.
48. Nield J., Balsera, M., De La Rivas, J., and Barber, J. (2002) *J. Biol. Chem. 277*, 15006−15012.
49. Tucker, D. L., Hirsch, K., Li, H., Boardman, B., and Sherman, L. A. (2001) *Biochim. Biophys. Acta 1504*, 409−422.
50. Shen, J.-R., and Inoue, Y. (1993) *Biochemistry 32*, 1825−1835.
51. Shen, J.-R., Qian, M., Inoue, Y., and Burnap, R. L. (1998) *Biochemistry 37*, 1551−1558.
52. Mayfield, S. P., Bennoun, P., and Rochaix, J.-D. (1987) *EMBO J. 6*, 313−318.
53. Burnap, R. L., and Sherman, L. A. (1991) *Biochemistry 30*, 440−446.

BI020228U