Defining the Bilin Lyase Domain: Lessons from the Extended Phytochrome Superfamily†

Shu-Hsing Wu‡ and J. Clark Lagarias*

Section of Molecular and Cellular Biology, University of California, Davis, Davis, California 95616

Received May 17, 2000; Revised Manuscript Received August 24, 2000

ABSTRACT: Through pattern searches of genomic databases, new members of the growing family of phytochrome-related genes were identified and used to construct a 130−180 amino acid motif that delimits the bilin lyase domain, a subdomain of the extended phytochrome family that is sufficient for covalent attachment of linear tetrapyrroles (bilins). To test this hypothesis, portions of locus *sll0821*, a novel phytochrome-related gene from *Synechocystis* sp. PCC6803 that encodes a large protein with two potential bilin binding sites, were amplified, and the recombinant apoproteins were tested for bilin binding and phytochrome photoactivity. Our experiments indicated that both sites of this protein, termed Cph2 for *c*yanobacterial *ph*ytochrome 2, possessed bilin lyase activity, revealing two distinct classes of bilin lyase domains—those whose bilin adducts are red, far-red reversible and a second class whose bilin adducts are nonphotochromic. Spectroscopic analysis of photochromic phycocyanobilin and fluorescent phycoerythrobilin adducts of a 24-kDa fragment of Cph2 definitively established that the motif identified by pattern searches represents a bona fide bilin lyase domain. Site-directed mutagenesis of highly conserved charged residues within bilin lyase domains of nearly all members of the extended phytochrome superfamily has identified a glutamate residue critical for bilin binding.

Phytochromes are biliprotein photoreceptors that mediate numerous adaptive responses of plants to their light environment (*1*, *2*). Light perception by phytochromes entails double-bond photoisomerization of the covalently linked linear tetrapyrrole (bilin) prosthetic group, which leads to a change in conformation and biochemical function of the associated protein moiety (*3*, *4*). The reversibility of this photoreaction is a hallmark of the phytochrome family, all members of which can interconvert between red light absorbing (Pr)[1] and far-red light absorbing (Pfr) forms. This property distinguishes phytochromes from the intensely fluorescent phycobiliproteins with similarly linked bilin prosthetic groups that function as photosynthetic antennae (*5*).

The characteristic photochemical and spectroscopic properties of phytochromes reflect the unique chemical environment of the bilin chromophore within their highly conserved 'photosensory' domain at the protein's N-terminus (see Figure 1). Roughly 500−600 amino acids (aa) in length, phytochrome photosensory domains are readily identifiable by their sequence similarity (i.e., >50% identity), while the C-terminal 'regulatory' domains of phytochromes are considerably more diverged. This similarity has facilitated identification of new phytochrome-related genes, such as the cyanobacterial phytochrome gene *cph1* (*6*, *7*), the *cph1*-related *bphP* genes from *Deinococcus radiodurans* and *Pseudomonas aeruginosa* (*8*), the phytochrome-related genes *plpA* and locus *sll0821* from *Synechocystis* sp. PCC6803 (*9*, *10*), *rcaE* from *Fremyella diplosiphon* (*11*), *ppr* from *Rhodospirillum centenum* (*12*), and *cikA* from *Synechococcus* sp. PCC7942 (*13*). Like plant apophytochromes, Cph1 and drBphP apoproteins yield red, far-red photoreversible photoreceptors upon incubation with bilins (*6*−*8*). Bilin binding to the products of *rcaE, ppr, cikA*, and *plpA* genes has not yet been described; however, genetic analyses show that all four genes encode proteins involved in light signaling (*9*, *11*−*13*).

The biochemical mechanism of bilin attachment to apophytochrome has been of interest to our lab for many years. On the basis of initial studies with apophytochrome from tetrapyrrole-deficient oat seedlings (*14*), it has been well established that apophytochromes are bilin lyases that catalyze thioether linkage formation with ethylidene-containing bilin chromophore precursors (for a review, see ref *15*). Indeed, the ability to form covalent adducts with phytochromobilin (PΦB) as well as the phycobilin analogues phycocyanobilin (PCB) and phycoerythrobilin (PEB) appears to be a property of all phytochromes. That the photosensory domain of phytochrome is wholly responsible for this bilin lyase activity has been well established by deletion mutagenesis, expression, and reconstitution of recombinant phytochromes (*16*−*18*). While the smallest region of this domain that exhibited bilin lyase activity was only 400 aa in length,

* Corresponding author. Phone: (530)752-1865. Fax: (530)752-3085.

† This work was supported by National Science Foundation Research Grant MCB 96-04511 to J.C.L. and by NIH National Center for Research Resources Award 1 P41 RR06009 to the Pittsburgh Supercomputing Center for computational biology support.

‡ Present address: Department of Plant Biology, Carnegie Institution of Washington, 260 Panama St., Stanford, CA 94305.

[1] Abbreviations: Cph1, *c*yanobacterial *ph*ytochrome 1; Cph2, *c*yanobacterial *ph*ytochrome 2; Cpl, *c*yanobacterial *p*hytochrome-*l*ike proteins; Pr, red light absorbing form of phytochrome; Pfr, far-red light absorbing form of phytochrome; PCB, phycocyanobilin; PΦB, phytochromobilin; PEB, phycoerythrobilin.

10.1021/bi001123z CCC: $19.00
Published on Web 10/13/2000

expression of phytochrome deletion mutants in planta suggests that a domain as small as 350 aa is capable of bilin binding (reviewed in ref *19*). A recent study showing that a 290 aa region of Cph1 is capable of forming photoactive bilin adducts in vitro corroborates this conclusion (*44*).

Because of the high conservation of phytochrome photosensory domains, it has been difficult to predict and identify residues critical for bilin attachment (*20*−*22*). As an alternative to site-directed mutagenesis of the very large number of residues conserved in the photosensory domains of known phytochromes, we exploited the expanding genomic databases to identify new phytochrome-related genes enabling us to more precisely delimit phytochrome's bilin lyase domain. Through expression of portions of a novel phytochrome-related gene from *Synechocystis* sp. PCC6803, our studies biochemically define two types of bilin lyase domain roughly 130−180 aa in length—one that yields red, far-red photoreversible phytochromes and a more diverged bilin lyase domain that yields photochemically inactive bilin adducts. The more limited number of conserved residues within lyase domains of the extended phytochrome superfamily has facilitated identification of a glutamate residue critical to bilin binding.

MATERIALS AND METHODS

Bioinformatics. Protein and nucleic acid database searches were performed using programs at publically available websites or by programs supported by the University of Wisconsin Genetics Computing Group (GCG) and by the Pittsburgh Supercomputer Biomedical Computing Group. The latest editions of the appropriate genomic databases were analyzed with the BLASTP and TBLASTN algorithms (*23*) at the NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) and Kazusa Research Institute Cyanobase (http://www.kazusa.or.jp/cyano/cyano.html) websites (*24*) or with the PROFILESEARCH (*25*) algorithm at the San Diego Supercomputer SEQWEB (http://www.sdsc.edu/projects/profile/new/html-auth.cgi) website. Motif searches were performed using an HMM algorithm at the PFAM (http://pfam.wustl.edu/) website or at the Molecular Pattern Website (http://www.sdsc.edu/mpr/) with the MEME HMM algorithm (*26*). Multiple sequence alignments were performed interactively at the Pittsburgh Supercomputing Center using the programs GCG PILEUP, SAGA (*27*), and CLUSTALW (*28*). In most cases, default scoring matrixes were utilized. Multiple sequence alignments were edited and annotated using SEQVU (*29*) or GENEDOC (http://www.psc.edu/biomed/genedoc).

Cloning of Fragments from Cyanobacterial Phytochromes 1 (slr0473) and 2 (sll0821). All PCR reactions were performed using Pfu-cloned DNA polymerase (Stratagene) and 50 ng of *Synechocystis* sp. PCC6803 genomic DNA with 35 cycles of 95 °C for 1 min, 55 °C for 2 min, and 72 °C for 3 min. Sense primers were designed with an *Xba*I site, optimized Shine−Dalgarno ribosome binding sites, and an ATG initiation codon; antisense primers were designed with a *Sal*I site for in-frame cloning to the Biometra pASK75B Strep-Tag expression vector (*30*) for optimal expression of the fusion protein. Primers Cph2P1S/*Xba*I, 5′-GC*TCTAGA*TAACGAGGGCAAAAAATGAACCCTAATCGATCCT-3′, and Cph2P390AS/*Sal*I, 5′-GC*GTCGAC*CTTCTTCCCAAGCCTCGAAG-3′ (restriction sites italicized), were designed for PCR amplification of a 1.1-kbp DNA fragment that corresponds to the N-terminal 390 aa of Cyanobase locus *sll0871*. We have named this clone Cph2−N390 for the N-terminal 390 aa region of cyanobacterial phytochrome 2. pCph2-N390ST was constructed by cloning the 1.1 kbp *Xba*I−*Sal*I restricted fragment into similarly digested pASK75B (Biometra) to create an in-frame fusion of Cph2−N390 with the Strep-Tag. Primer pair Cph2P1S and Cph2P197AS/*Sal*I, 5′-GC*GTCGAC*CCTGATGTACCTGCCGGCTTAA-3′, and primer pair Cph2P853S/*Xba*I, 5′-GC*TCTAGA*TAACGAGGGCAAAAAATGACCTATCTTTACTATC-3′, and Cph2P1276AS/*Sal*I, 5′-GC*GTCGAC*CAACTTCCCCATCAACATGG-3′, were similarly used for the amplification of 600-bp and 1.27-kbp DNA fragments encoding the N-terminal 197 aa region of Cph2 (i.e., Cph2−N197) and the C-terminal 423 aa region of Cph2 (i.e., Cph2−C423), respectively. After *Xba*I and *Sal*I digestions, these fragments were cloned into pASK75B as described above. PCR-based mutagenesis (*31*) was employed for site-specific mutagenesis of Cph1−N514 clone in pASK75B vector described in a previous paper (*32*). Sense and antisense primers, Cph1P1S/*Spe*I, 5′-GC*ACTAGT*TAACGAGGGCAAAAAATGGCCACCACCGTAC-3′, and Cph1P514AS/*Sal*I, 5′-GC*GTCGAC*CACCTTCTTCTGCCTGGCGCAA-3′, were used to amplify the 1.5-kbp region of Cph1 (Cyanobase locus *slr0473*) for cloning in pASK75B as described above. Mutagenesis primers including DR171-2, 5′-TTCATCAAAGCGGTCCAGCATCACG(G/C)CG(G/C)CAAAGCCAGTCATACG-3′, R172AG/*Rsa*I, 5′-ATGACTGGCTTTGACG(G/C)GGTGATGCT*GTAC*CGCTTTGATGAA-3′, E189N/*Bst*UI, 5′-GGGTTCCATATCAT*CGCG*TTTATCNNNGGCAATGACATCACC-3′, R222G/*Bsa*HI, 5′-ATCCACAACCCCATT*GGCGTC*ATTCCCGATGTTTAT-3′, and D226G/*Bsa*HI, 5′-ATTCGAGTAATTCCC*GGCGTC*TATGGTGTGGCGGTG-3′, were used to amplify Cph1−N514 with amino acids substitutions at the residue number indicated. The *Spe*I−*Sal*I restricted PCR fragment was cloned into *Xba*I−*Sal*I restricted pASK75B for the expression of strep-tagged versions of mutant Cph1−N514. A restriction enzyme recognition site was designed as a silent mutation in each mutagenesis primer that enabled identification of the mutant clones by restriction enzyme digestion. The DNA sequences of all clones were confirmed.

Expression of Recombinant Apophytochromes in E. coli. A 30-mL overnight culture of *E. coli* strain DH5α harboring the plasmid for each specific construct was used to inoculate 1 L of Luria broth medium supplemented with 100 μg/mL ampicillin. The culture was incubated at 37 °C with shaking until the OD_{550} reached 0.5, at which time 0.2 μg/mL anhydrotetracycline (final concentration) was added to the cultures to induce recombinant protein expression. Following induction at 30 °C for 3 h, cell pellets were collected by centrifugation at 3000*g* for 10 min and washed once with wash buffer (20 mM Tris-HCl, pH 8.0, 20 mM NaCl, and 1 mM EDTA). Washed cell pellets were resuspended in 15 mL of lysis buffer (50 mM Tris-HCl, pH 8.0, at 4 °C, 100 mM NaCl, 0.05% NP40, 2.0 μg/mL leupeptin, 2 mM benzamidine, 2 mM PMSF, 3 μg/mL pepstatin A, and 1mM DTT) and frozen in liquid N_2. Thawed cells were lysed with a single pass through a French Press set at 10 000 psi, and the cell lysate was ultracentrifuged at 100000*g* for 30 min. The soluble protein supernatant was further fractionated by

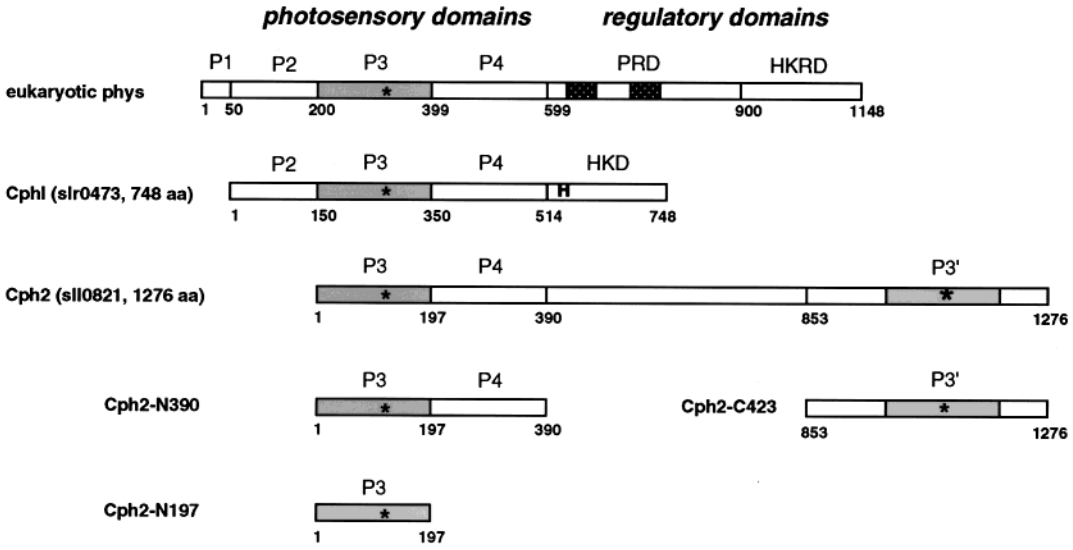

FIGURE 1: Domain structure of eukaryotic phytochromes, Cph1, Cph2, and Cph2 deletion constructs used in this study. The photosensory domains, P1−P4, correspond to subdomains defined by bioinformatics, natural deletion mutagenesis, and work presented in this study. PRD and HK(R)D represent PAS-related domains and histidine kinase-related domains (see ref *43* for discussion). The two direct repeats within the PRD are shown as dark dotted boxes. Asterisks mark the conserved cysteine residue found within each of the two types of bilin lyase GAF domains defined by this work, P3 and P3′ (shaded).

adding 0.23 g/mL ground, ultrapure ammonium sulfate, incubating on ice at least 1 h and centrifuging at 17000*g* for 30 min at 4 °C. The resulting protein precipitate was then redissolved in 10 mL of buffer W (100 mM Tris-HCl, pH 8.0, containing 1 mM EDTA, 1 mM DTT, and 1 mM PMSF), clarified by centrifugation for 15 min at 38000*g*, and used for bilin assembly, difference spectral analyses, and/or affinity chromatography.

Holophytochrome Reconstitution, Absorbance, and Fluorescence Measurements. For in vitro reconstitution of holophytochrome, 3E-phycocyanobilin (PCB) or phycoerythrobilin (PEB) was added to the apophytochrome-containing extract to give a final bilin concentration of 4−10 μM, and the mixture was incubated at room temperature under dim green safe light for 30 min prior to spectrophotometric measurements. Absorbance spectra and absorbance difference measurements were determined with HP8450A or HP8453 UV/visible spectrophotometers using the protocol described previously (*33*). Fluorescence excitation and emission spectra of the PEB adduct of apophytochromes were determined with an SLM Aminco Bowman AB2 fluorimeter (*32*).

Affinity Purification of Recombinant Strep-Tagged Proteins. The recombinant proteins from *E. coli* were purified by the protocol described previously (*30*, *34*). Streptavidin (core protein)−Sepharose (2.5−3 mL bed volume) was packed in columns and washed with 5−10 bed volume of buffer W prior to use. Avidin-blocked, clarified ammonium sulfate-fractionated protein was applied to the column, and the column was then washed by 15−20 bed volume of buffer W (see above) to remove the nonspecifically bound proteins. Strep-tagged protein was eluted with 5 bed volume of buffer E (100 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 3 mM diaminobiotin). Eluted protein was concentrated and buffer exchanged to TEGE buffer (25 mM Tris-HCl, pH 8.0, 25% ethylene glycol, and 1 mM EDTA) using an Ultrafree-15 centrifugal filter device (Millipore).

SDS−PAGE, Zinc-Blot, and Immunoblot Analyses. Protein samples were analyzed by SDS−PAGE with the Laemmli buffer system (*35*). After electrophoresis, proteins were electrophoretically transferred to poly(vinylidene difluoride) (PVDF) membranes (Immobilon P, Millipore) for 1 h at 100 V. The membrane was then used for zinc-blot and immunoblot analyses. Zinc-blot analysis was performed as described previously utilizing a Molecular Dynamics Storm 860 instrument with setting of red-fluorescence and PMT = 1000 (*36*). Figures 3 and 6 are shown as inverted images (i.e., dark bands represent fluorescent bands). For immunoblot detection, recombinant phytochrome was probed with alkaline phosphatase conjugated streptavidin (Amersham) and developed as described previously (*30*).

RESULTS

Database Searches Identify New Members of the Phytochrome Superfamily. To identify new phytochrome-related sequences, the highly conserved region surrounding the bilin binding cysteine residue of Cph1 was used as a query sequence for BLASTP searches of the *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC7120 genome databases. These searches identified a large family of phytochrome-related genes in both genomes, i.e., 12 and 13 genes from *Synechocystis* and *Anabaena*, respectively, whose protein products shared one to four repeats of a common motif, roughly 150−200 aa in length (indicated schematically in Figure 1 as domain P3). Previously identified as a GAF domain, a conserved 'domain' that is found on a diverse group of phototransducing proteins (*37*), the phytochrome-related protein motif identified by our studies more narrowly defines a subclass of GAF domain-containing proteins.

A weighted profile of this motif was then used to find additional members of the extended phytochrome family in the publically available protein and nucleic acid databases. These included the phytochrome-related genes *rcaE* from *Fremyella diplosiphon* (*11*), *ppr* from *Rhodospirillum centenum* (*12*), *cikA* from *Synechococcus* sp. PCC7942 (*13*), and the *bphP* genes from *Deinococcus radiodurans* and *Pseudomonas aeruginosa* (*8*). The phytochrome-related GAF motif was

further delimited by the identification of genes whose putative protein products either initiated or terminated at the N- or C-termini of this motif. Two of these proteins, yph1_synp2 from the cyanobacterium *Synechococcus* sp. PCC7002 (GB:P32039) (*38*) and an ORF from the *Anabaena* sp. PCC7120 BAC contig *c200*, are 153 and 179 aa in length, which supports the hypothesis that this motif corresponds to a bona fide protein domain. An optimized multiple sequence alignment of this region, including repeats, was then constructed using the programs CLUSTALW, SAGA, and MEME (Figure 2). Although pairwise sequence identities were as low as 7%, profile analysis using this alignment yielded statistically significant scores for all proteins used to construct the alignment (data not shown). On the basis of these pattern recognition analyses and data provided in this manuscript, we assert that the 130−180 residue motif defines a bilin lyase domain.

Locus sll0821 Encodes a New Phytochrome with Two Functional Bilin Binding Sites. Phylogenetic tree reconstruction using this alignment indicates that the GAF domains of Cph1, the phytochrome-related ORF within *Anabaena* locus *c371*, both Bphs, and AphB—a Cph1-related protein from *Anabaena* (GenBank Accession No. AB034952)—are the members of the phytochrome family most closely related to eukaryotic phytochromes (data not shown). Indeed, two members of this Cph1 subfamily, notably Cph1 and drBphP, have already been confirmed to be phytochromes by their ability to assemble with bilins to produce photoactive biliproteins (*6*−*8*). The next most closely related loci to plant phytochromes are *Synechocystis* locus *sll0821* (first noted by Manabe (*10*)) and an ORF in *Anabaena* locus *c297a*. Both loci encode large polypeptides (i.e., 1276 and 1286 aa, respectively), each with two GAF domains. The N-terminal GAF domains of both proteins are more similar to phytochromes (i.e., 23−29% identity), whereas the C-terminal domains are more diverged (i.e., 18−25% identity). Like the vast majority of the other phytochrome-related loci shown in Figure 2, the C-terminal GAF domains of ORF *sll0821* and ORF *c297a* lack a 25 aa region near the putative bilin attachment cysteine residue (i.e., between residues 113 and 138 in Figure 2). On the basis of this similarity, the data presented below, and a recent study on this locus which appeared during review of this work (*39*), the product of locus *sll0821* has been named *c*yanobacterial *ph*ytochrome 2 or Cph2.

To test whether the *cph2* locus encodes a phytochrome, it was amplified as a full-length ORF and in two parts, corresponding to the N-terminal 390 aa residues (i.e., Cph2−N390) and to the C-terminal 423 aa residues (i.e., Cph2−C423). These proteins were expressed in *E. coli* as Strep-tagged recombinant proteins to facilitate their identification on blots and for affinity purification. To avoid the complication of the presence of two bilin sites in full-length Cph2, initial analyses focused on the two truncated proteins Cph2−N390 and Cph2−C423. Both apoproteins were highly expressed yielding polypeptides of the correct molecular masses (i.e., 45−50 kDa). To establish if the two apoproteins could covalently attach to bilins, crude soluble protein extracts were incubated with PCB. Unfortunately Cph2−N390 proved to be mostly insoluble; however, the small amount that remained soluble was further analyzed for bilin binding. Figure 3 shows that both Cph2−N390 and Cph2−C423 apoproteins yielded covalent adducts with PCB, as verified by the orange fluorescence observed for the respective apoprotein bands after incubation with zinc ions (*36*). Moreover, the PCB adduct of Cph2−N390 exhibited a characteristic red, far-red difference spectrum establishing that Cph2 is a bona fide phytochrome (Figure 4A). This was confirmed with full-length Cph2, which also yielded a PCB adduct with a very similar difference spectrum to that of Cph2−N390 (*39*; data not shown). By contrast, the PCB adduct of Cph2−C423 was photochemically silent (Figure 4A). In addition, neither Cph2−N390 nor Cph2−C423−PCB adducts exhibited significant fluorescence emission, suggesting that the excited states of both adducts were quenched by radiationless de-excitation processes (data not shown).

For comparative purposes, the difference spectra of the PCB adducts of recombinant preparations of oat phytochrome A (AsphyA), Cph1, and Cph2−N390 are depicted in Figure 4A. This comparison reveals that the Pr and Pfr absorption maxima of the PCB adduct of Cph2−N390 are similar to those of the other phytochromes, except for a blue-shift of 10−16 nm. This indicates that the structure and environment of the Cph2−N390 prosthetic group is qualitatively similar to those of other phytochromes, despite the wide divergence in the sequence of their photosensory domains. To more fully examine the optical properties of Cph2−C423, the assembled protein was affinity purified. In contrast with phytochromes, the PCB adduct of Cph2−C423 exhibited greater absorption at its blue absorption maximum at 421 nm than at its red absorption maximum at 634 nm (Figure 4B). This type of spectrum is more similar to those of porphyrins and/or metalloporphyrins, suggesting that the bilin adopts a more cyclohelical configuration when bound to the Cph2−C423 apoprotein (*40*). While the exact structures of the PCB adducts of Cph2(C390) and Cph2−C423 remain to be determined, these analyses clearly indicate that locus *cph2* encodes a bona fide phytochrome with two functional bilin lyase domains.

Phytochrome Bilin Lyase Activity Resides in a 200 aa Core Domain. Comparison of the domain organization of eukaryotic phytochromes, Cph1 and Cph2 (see Figure 1), indicates that the N-terminal 200 aa region of eukaryotic phytochromes is unnecessary for bilin binding and photoreversibility and is therefore dispensable. As has been well documented by many laboratories, removal of the N-terminal 50 aa region or P1 subdomain leads to a blue shift of the Pfr spectrum (see Figure 4A). That the P2 subdomain also is dispensable for bilin attachment is confirmed by the phytochrome-like properties of the PCB adduct of Cph2−N390, which lacks this subdomain (see Figure 1). This information together with the observation that the N-terminal 399 aa region of oat phytochrome can functionally assemble with bilins when expressed in transgenic tobacco plants (*41*) suggest that the N-terminal 200 residues of Cph2 should be sufficient for bilin binding. To test this hypothesis, we expressed a polypeptide containing only the N-terminal 197 aa region of Cph2 in *E. coli* and incubated the soluble protein extract with the bilins, PCB, or PEB. That the 24-kDa recombinant Cph2−N197 protein yielded covalent adducts with both bilins was shown by zinc-blot analyses (Figure 5A). This result establishes that the P3 domain of Cph2 is sufficient for bilin attachment and is therefore a true bilin lyase domain.

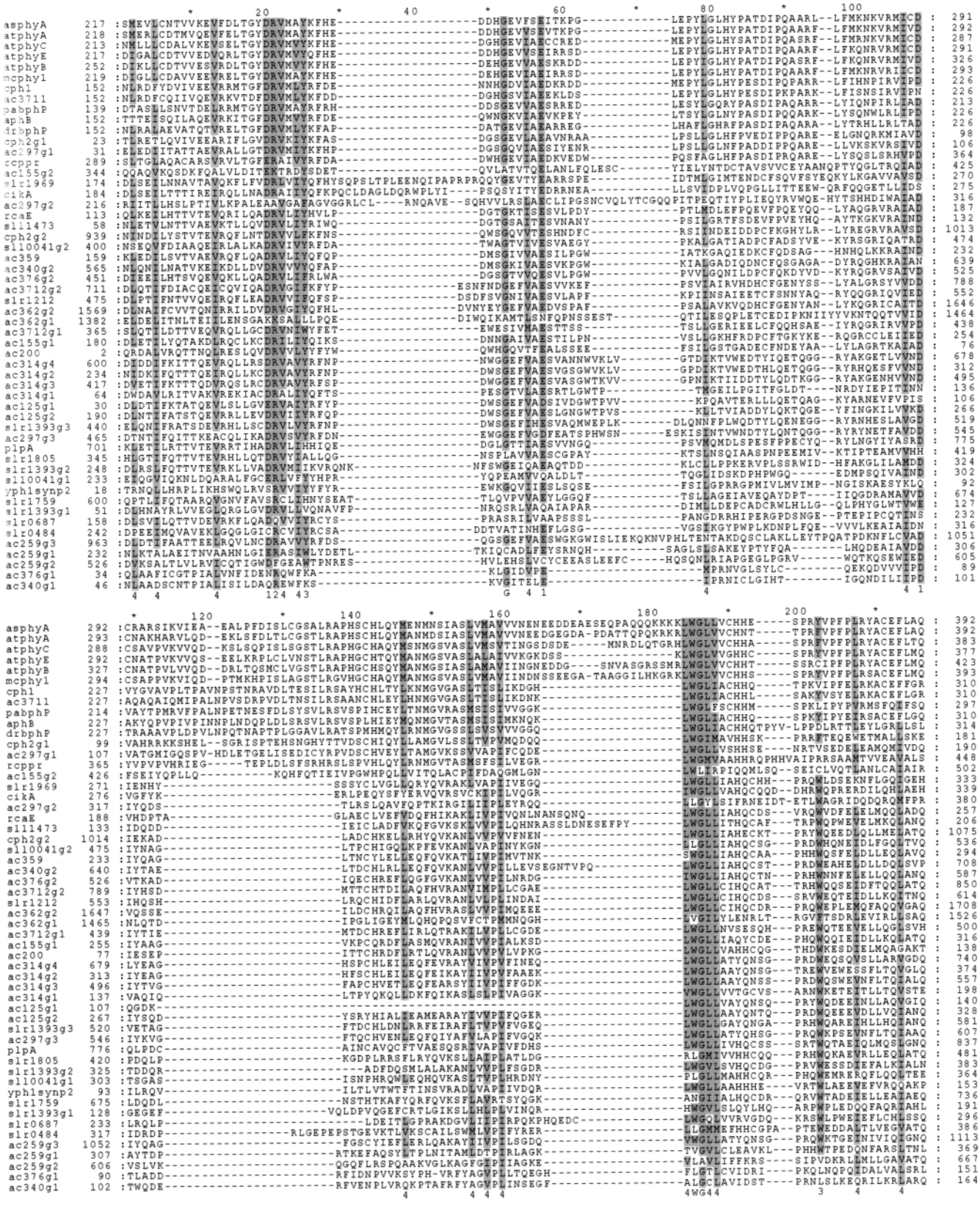

FIGURE 2: Multiple sequence alignment of bilin lyase GAF domains identified by pattern searches described in the text. Sequence similarity groups (1, D = E; 2, R = K; 3, F = Y = W; and 4, L = I = V = M), shown in the consensus sequence reflect conservation in more than 75% of the sequences and are shaded. Sequence IDs reflect the gene name, Cyanobase loci for *Synechocystis* sp. PCC 6803 genes (indicated with the prefix 'sll' or 'slr'), or contig locations for *Anabaena* sp. PCC 7120 loci (indicated with the prefix 'ac'). The gene ID suffixes, g1−g4, reflect multiple bilin lyase GAF domains in a given protein. Residue numbers for each protein sequence are indicated. Database accession numbers are SP:P06593 for asphyA, SP:P14712 for atphyA, SP:P14713 for atphyB, SP:P14714 for atphyC, SP:P42498 for atphyE, GB:U31284 for mcphy1, SP:Q55168 for cph1, DBJ:BAA10536 for cph2/sll0821, GB:U59741 for rcaE, GB:U67397 for plpA, GB:AAF12261.1 for drbphP, GNL:PAGP_287 for pabphP, GB:AAD22391.1 for rcppr, GB:AF258464 for cikA, and GB:P32039 for yph1synp2.

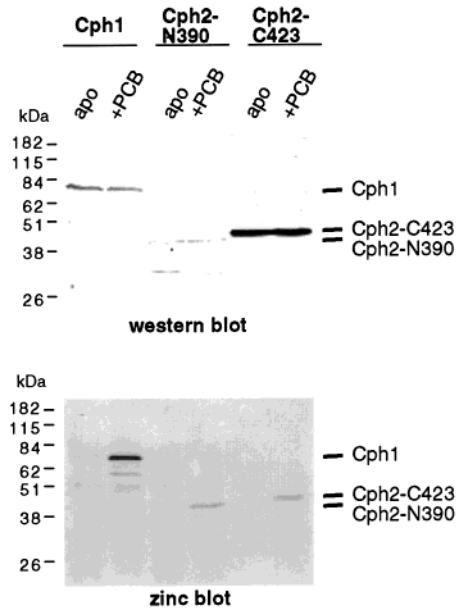

FIGURE 3: Expression and assembly of Cph2−N390 and Cph2−C423. *E. coli* strains harboring plasmid encoding Cph1, Cph2−N390, or Cph2−C423 were induced to express recombinant proteins as described in Materials and Methods. Soluble protein extracts were fractionated on SDS−PAGE followed by immunoblot (upper panel) and zinc-blot analyses (lower panel). apo, apoprotein only; +PCB, 4 μM PCB was preincubated with soluble protein extract prior to gel electrophoresis analysis.

To allow further investigation of its spectroscopic properties, Cph2−N197 was affinity purified. As shown in Figure 5B, the PCB adduct of Cph2−N197 was photochromic although the difference spectrum was markedly different from those of other phytochromes. In comparison to the PCB adduct of the photosensory domain of Cph1 (i.e., Cph1−N514), Cph2−N197 displayed a blue-shifted Pfr maximum (i.e., 690 nm vs 706 nm) and reduced the amount of conversion to the Pfr form (Figure 5B). Cph2−N197 also exhibited a broader absorption peak and a smaller ratio of the red-to-blue light absorption maxima compared to those of the PCB adduct of Cph1. Like PEB adducts of other phytochromes, the PEB adduct of Cph2−N197 was strongly fluorescent (Figure 5C), with fluorescence excitation and emission spectra very similar to those of PEB adducts of apophytochromes, known as phytofluors (*32*). This indicates that PEB is bound to Cph2−N197 in a configuration very similar to that of other phytofluors.

Functional Analysis of Conserved Charged Residues in the Bilin Lyase Domain. The alignment of the phytochrome-related protein family shown in Figure 2 reveals that the most highly conserved residues of the lyase domain possess hydrophobic side chains with only four charged residues, i.e., aspartate-20, arginine-21, glutamate-57, and aspartate-108, being conserved in more than 80% of the proteins. Interestingly, the site of bilin attachment, i.e., cysteine-141, was preserved in only 53−55% of the aligned proteins. To test the role of these charged residues in bilin lyase activity, site-directed mutations were introduced into Cph1−N514, and the effects of these mutations on bilin binding and photoactivity were examined. As shown in Figure 6, aspartate-20 (i.e., D171 on Cph1) and arginine-21 (i.e., R172 on Cph1) could be substituted with alanine or glycine without affecting the binding of PCB. Such mutations also had no

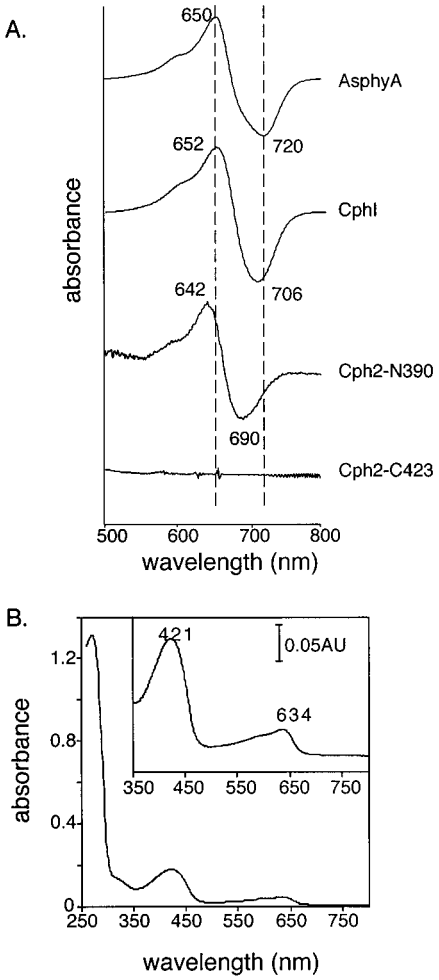

FIGURE 4: Spectrophotometric analyses of recombinant Cph2 bilin lyase domains. (A) Difference spectra of PCB adducts of Cph2−N390 and Cph2−C423 adducts were measured as described in Materials and Methods. Difference spectra of recombinant AsphyA and Cph1 are presented for comparison. Absorption maxima and minima are indicated in nm. (B) Absorption spectrum of affinity-purified Cph2−C423−PCB adduct. The absorption between 350 and 800 nm is rescaled and shown in the insert with the absorption unit (AU) indicated.

effect on the difference spectra. Only by converting both residues to alanine (or glycine) was bilin binding abolished. In contrast, all site-directed mutations tested for glutamate 57 (i.e., E189 on Cph1)—which included substitutions with alanine, glycine, lysine, and threonine—abolished bilin attachment. These results show that glutamate 57 (i.e., E189) is essential for catalysis, but individually, aspartate-20 and arginine-21 are not critical for bilin attachment and holoprotein photoreversibility.

DISCUSSION

These investigations document the predictive utility of examining the protein and nucleic acid databases to delimit a subdomain in phytochrome that is sufficient for one of its biochemical activities (i.e., bilin binding). By taking advantage of the deletion mutagenesis already performed by nature, we have made progress in defining the borders of this functional subdomain that would have been difficult to define by directed mutagenesis. Indeed, that the N-terminal 197 aa of Cph2 combines with PCB to yield a photoactive holophytochrome convincingly establishes that the P3 GAF

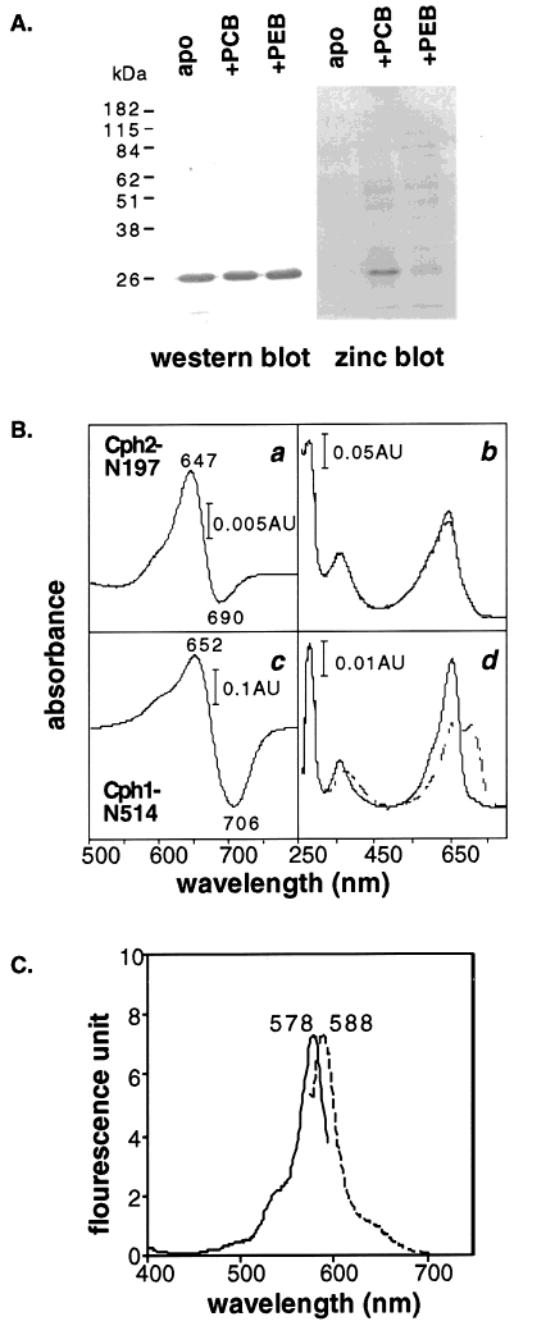

FIGURE 5: Biochemical and spectral analyses of recombinant Cph2−N197. (A) Soluble protein extracts containing recombinant Cph2−N197 were fractionated on SDS−PAGE followed by immunoblot (left panel) and zinc-blot analyses (right panel). apo, apoprotein only; +PCB/+PEB, 4−10 μM PCB/PEB was preincubated with soluble protein extract prior to gel electrophoresis analysis. (B) Difference and absorption spectra of Cph2−N197−PCB adduct (a, b) were constructed and compared with the ones of Cph1−N514−PCB adduct (c, d). Absorption maxima and minima are indicated in nm. (C) Fluorescence excitation (solid line) and emission (dashed line) spectra of Cph2−N197−PEB adduct.

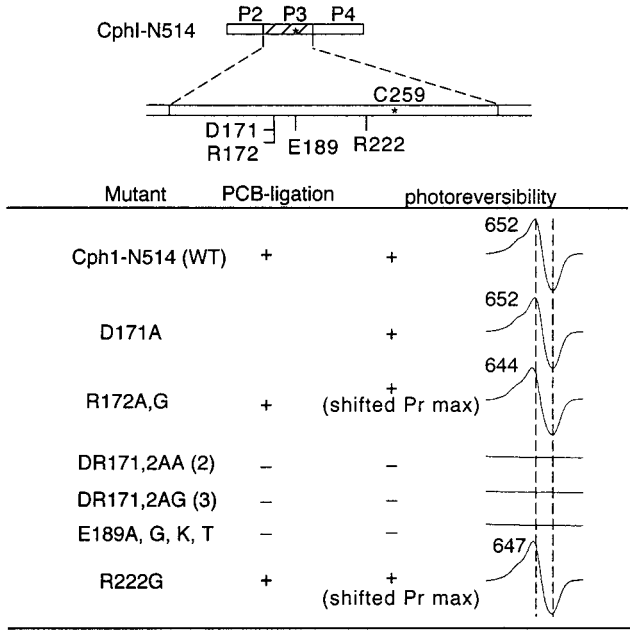

FIGURE 6: Lyase activity and spectral properties of Cph1−N514 mutants. (A) Conserved charged residues in chromophore lyase domain selected for site-directed mutagenesis. C259 (*) is the putative chromophore attachment site in Cph1. (B) Lyase activity of Cph1−N514 and each mutant were determined by the presence (+) or absence (−) of fluorescence on zinc-blot analysis as described in Materials and Methods. Difference spectrum of each construct was also measured. Abbreviations used: D, aspartate; A, alanine; R, arginine; G, glycine; E, glutamate; K, lysine; T, threonine.

domain defined by the alignment in Figure 2 represents a bona fide bilin lyase domain.

Spectroscopic data presented in this report also indicate that the P3 GAF domain is sufficient for tethering the bilin A-, B-, and C-ring pyrrole rings in a configuration similar to that of the native phytochrome chromophore. Indeed, the striking similarity of the fluorescence excitation and emission spectra of the PEB adduct of Cph2−N197 with those of known phytochromes, aka the phytofluors (*32*), strongly supports this interpretation. The reduced photoconversion efficiency, broader absorption envelope, and reduced ratio of the red-to-blue light absorption maxima of the PCB adduct of Cph2−N197 are consistent with those of a biliprotein whose bilin prosthetic group is less extended and less rigidly held than those of native phytochromes. Since a more 'typical' phytochrome difference spectrum is restored when the adjacent P4 domain is present as was seen for the PCB adduct of Cph2−N390, the P4 domain appears to serve a dual role—both to stabilize the extended configuration of the Pr chromophore and to stabilize the Pfr form of the bilin prosthetic group.

Our studies also suggest that molecular evolution of a phytochrome progenitor yielded a second class of phytochrome-related proteins in cyanobacteria with bilin lyase domains that bind bilins to afford nonphotoactive biliproteins. Typified by the C-terminal bilin lyase domain on Cph2, we propose that other members of this *c*yanobacterial *p*hytochrome-*l*ike (Cpl) family are biliproteins whose bilin prosthetic groups adopt a more cyclic, porphyrin-like configuration. In this regard, we have observed that PCB covalently binds to recombinant Slr1969, another member of the cpl family. Like Cph2−C423, the PCB adduct of Slr1969 is not red, far-red reversible; absorbs light more strongly in the blue region; and is nonfluorescent (data not shown). These spectrophotometric properties are consistent with a porphyrin-like configuration of their bilin prosthetic groups, in which the excited state is quenched by phototautomerization of adjacent pyrrole rings that share an intramolecular H-bond (*40*). On the basis of the multiple sequence alignment of the bilin lyase domains of the extended phytochrome superfam-

ily, we propose that the presence of the 20−23 aa region adjacent to the putative cysteine binding site (i.e., residues 114−138 in Figure 2), distinguishes the phytochrome photoreceptor family from the cpl family. The functional significance of this region of phytochromes is further underscored by the loss of function L309S atphyA mutation (C. Fairchild and P. H. Quail, personal communication) and S349F atphyB mutation (*42*), which both map to this region.

With regard to the nature of the bilin−apoprotein association in the Cpls, it is important to note that the conserved cysteine residue (i.e., cys141 in Figure 2) is not present on all of the bilin lyase domains represented in Figure 2. While it is conceivable that many of these phytochrome-related domains might bind bilins through a different residue(s), such as the histidine adjacent to cys141 as was reported for drBphP (*8*), we favor the conserved cysteine to be the site of bilin attachment for all phytochrome-related proteins that possess this residue. This includes both lyase domains of Cph2 and also Slr1969, whose single bilin lyase domain lacks the adjacent histidine residue. Indeed, recent mutagenesis experiments with cph2 strongly implicate this conserved cysteine in the P3 domain to be the site of bilin attachment (*39*). Nevertheless, a large number of the bilin lyase domains depicted in Figure 2 lack the conserved cysteine residue (i.e., 46% of the bilin lyase domains shown in Figure 2). It is possible that these domains noncovalently interact with bilins, or alternatively that they interact with other tetrapyrrole ligands such as hemes or porphyrins, to allosterically regulate the activity of an associated functional domain(s).

A multiple sequence alignment of the lyase superfamily has also revealed a limited number of conserved charged residues that we have begun to mutagenize to test their potential role in bilin attachment, photoreversibility, and transduction of the light signal. In this regard, we have identified a conserved glutamate residue that is critical for bilin attachment. It is likely that this residue will be involved either in bilin binding and/or catalysis. Utilizing saturation mutagenesis and chemically altered bilins, we ultimately hope to elucidate the chemical role of this glutamate residue in the mechanism of bilin attachment to apophytochrome. Among the possibilities include protonation of the B- or C-ring pyrrole nitrogens, a role as a general base catalyst to increase the nucleophilicity of the cysteine residue, or its ability to enhance the electrophilicity of the bilin precursor via H-bonding with either a pyrrolic nitrogen or a lactam carbonyl oxygen.

The most surprising result from our mutagenesis studies was the observation that the highly conserved aspartate-20 and arginine-21 residues were not critical for bilin binding in mutants that lack one or the other residue. Indeed, since the bilin prosthetic group possesses two propionate moieties, we expected that the protonated guanidium side chain of the conserved arginine-21 residue might be involved in anchoring the bilin to the apoprotein via a salt linkage. While our data appear to rule out a direct role of this arginine residue in bilin binding, the loss of function of the asp20/arg21 double mutant indicates that these residues are important contributors to the proper folding and/or overall conformation of the bilin binding site of the lyase domain. We also examined the role of the next most conserved arginine in the lyase subdomain, arginine-106 in the alignment in Figure 2 that corresponds to arginine-222 in Cph1, via its mutagenesis to glycine. Figure 6 shows that this mutant still bound bilin and yielded a photoactive holophytochrome. The possibility that other less conserved basic residues play a role in propionate binding (i.e., arginine-95, arginine-210, or arginine/lysine-203), perhaps in bona fide phytochromes only, remains to be addressed in future investigations. We would like to know the functional role(s) of aspartate-21 and arginine-22 that are conserved in more than 80% of the phytochrome superrfamily. One intriguing possibility is that these residues play a universal role in transducing the light and/or ligand binding signal by the phytochromes, Cphs, Cpls, and possibly other GAF domain-containing proteins to other associated functional domains in these proteins. The role that these residues play in the light-dependent regulation of protein kinase activity of phytochromes (*7*, *43*) will be the subject of a future investigation.

ACKNOWLEDGMENT

We gratefully acknowledge Lawrence Eng and Loi Lam for construct preparation and experimental verification. We thank Nicole Frankenberg, Greg Gambetta, and Beronda Montgomery for critical reading of this manuscript. We thank Drs. Susan S. Golden, Craig Fairchild, and Peter Quail for communication of unpublished data.

REFERENCES

1. Furuya, M. (1993) *Annu. Rev. Plant Physiol. Plant Mol. Biol. 44*, 617−645.
2. Quail, P. H. (1997) *Plant Cell Environ. 20*, 657−665.
3. Braslavsky, S. E., Gärtner, W., and Schaffner, K. (1997) *Plant Cell Environ. 20*, 700−706.
4. Song, P. S. (1999) *J. Biochem. Mol. Biol. 32*, 215−225.
5. Glazer, A. N. (1994) *J. Appl. Phycol. 6*, 105−112.
6. Hughes, J., Lamparter, T., Mittmann, F., Hartmann, E., Gärtner, W., Wilde, A., and Börner, T. (1997) *Nature 386*, 663−663.
7. Yeh, K.-C., Wu, S.-H., Murphy, J. T., and Lagarias, J. C. (1997) *Science 277*, 1505−1508.
8. Davis, S. J., Vener, A. V., and Vierstra, R. D. (1999) *Science 286*, 2517−2520.
9. Wilde, A., Churin, Y., Schubert, H., and Börner, T. (1997) *FEBS Lett. 406*, 89−92.
10. Manabe, K., and Nakazawa, M. (1997) *J. Plant Res. 110*, U3−U3.
11. Kehoe, D. M., and Grossman, A. R. (1996) *Science 273*, 1409−1412.
12. Jiang, Z. Y., Swem, L. R., Rushing, B. G., Devanathan, S., Tollin, G., and Bauer, C. E. (1999) *Science 285*, 406−409.
13. Schmitz, O., Katayama, M., Williams, S. B., Kondo, T., and Golden, S. S. (2000) *Science 289*, 765−768.
14. Elich, T. D., and Lagarias, J. C. (1989) *J. Biol. Chem. 264*, 12902−12908.
15. Terry, M. J., Wahleithner, J. A., and Lagarias, J. C. (1993) *Arch. Biochem. Biophys. 306*, 1−15.
16. Deforce, L., Tomizawa, K. I., Ito, N., Farrens, D., Song, P. S., and Furuya, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10392−10396.
17. Hill, C., Gärtner, W., Towner, P., Braslavsky, S. E., and Schaffner, K. (1994) *Eur. J. Biochem. 223*, 69−77.
18. Gärtner, W., Hill, C., Worm, K., Braslavsky, S. E., and Schaffner, K. (1996) *Eur. J. Biochem. 236*, 978−983.
19. Cherry, J. R., and Vierstra, R. D. (1994) in *Photomorphogenesis in Plants* (Kendrick, R. E., and Kronenberg, G. H. M., Eds.) pp 271−297, Martinus Nijhoff Publishers, Dordrecht, The Netherlands.
20. Deforce, L., Furuya, M., and Song, P. S. (1993) *Biochemistry 32*, 14165−14172.

21. Bhoo, S. H., Hirano, T., Jeong, H. Y., Lee, J. G., Furuya, M., and Song, P. S. (1997) *J. Am. Chem. Soc. 119*, 11717−11718.
22. Remberg, A., Schmidt, P., Braslavsky, S. E., Gärtner, W., and Schaffner, K. (1999) *Eur. J. Biochem. 266*, 201−208.
23. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
24. Nakamura, Y., Kaneko, T., and Tabata, S. (2000) *Nucleic Acids Res. 28*, 72.
25. Gribskov, M., and Veretnik, S. (1996) *Methods Enzymol. 266*, 198−212.
26. Grundy, W. N., Bailey, T. L., Elkan, C. P., and Baker, M. E. (1997) *CABIOS, Comput. Appl. Biosci. 13*, 397−406.
27. Notredame, C., and Higgins, D. G. (1996) *Nucleic Acids Res. 24*, 1515−1524.
28. Higgins, D. G., Thompson, J. D., and Gibson, T. J. (1996) *Methods Enzymol. 266*, 383−402.
29. Garcher, J. (1995) SEQVU V.1.1, Garvan Institute of Medical Research, Sydney, Australia.
30. Murphy, J. T., and Lagarias, J. C. (1997) *Photochem. Photobiol. 65*, 750−758.
31. Picard, V., Ersdalbadju, E., Lu, A. Q., and Bock, S. C. (1994) *Nucleic Acids Res. 22*, 2587−2591.
32. Murphy, J. T., and Lagarias, J. C. (1997) *Curr. Biol. 7*, 870−876.
33. Litts, J. C., Kelly, J. M., and Lagarias, J. C. (1983) *J. Biol. Chem. 258*, 11025−11031.
34. Voss, S., and Skerra, A. (1997) *Protein Eng. 10*, 975−982.
35. Laemmli, U. K. (1970) *Nature 227*, 680−685.
36. Wu, S.-H., and Lagarias, J. C. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8989−8994.
37. Aravind, L., and Ponting, C. P. (1997) *Trends Biochem. Sci. 22*, 458−459.
38. Schluchter, W. M., and Bryant, D. A. (1992) *Biochemistry 31*, 3092−3102.
39. Park, C.-M., Kim, J.-I., Yang, S.-S., Kang, J.-G., Kang, J.-H., Shim, J.-Y., Chung, Y.-H., Park, Y.-M., and Song, P.-S. (2000) *Biochemistry 39*, 10840−10847.
40. Falk, H. (1989) *The Chemistry of Linear Oligopyrroles and Bile Pigments*, Springer-Verlag, Vienna.
41. Cherry, J. R., Hondred, D., Walker, J. M., Keller, J. M., Hershey, H. P., and Vierstra, R. D. (1993) *Plant Cell 5*, 565−575.
42. Bradley, J. M., Murphy, G. P., Whitelam, G. C., and Harberd, N. P. (1996) *J. Exp. Bot. 47*, 1449−1455.
43. Yeh, K.-C., and Lagarias, J. C. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13976−13981.
44. Park, C.-M., Shim, J.-Y., Yang, S.-S., Kang, J.-G., Kim, J.-I., Luka, Z., and Song, P.-S. (2000). *Biochemistry 39*, 6349−6356.

BI001123Z