# Sequence similarities of phytochrome to protein kinases: implication for the structure, function and evolution of the phytochrome gene family

Fritz Thümmler[a,*], Patricia Algarra[a], Gisela M. Fobo[b]

[a]*Botanisches Institut der Universität München, Menzingerstr. 67, D-80638 München, Germany*
[b]*Martinsried Institute for Protein Sequences (MIPS) am Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried bei München, Germany*

Received 22 October 1994; revised version receieved 21 November 1994

**Abstract Phytochrome, the best characterised plant photoreceptor, is encoded by a small multigene family within the plant kingdom. The different phytochrome types are composed of a conserved light-sensing chromophore domain of about 80 kDa and a less-conserved C-terminal domain of about 50 kDa. The C-terminus of phytochrome of the moss *Ceratodon purpureus* is homologous to the catalytic domain of eukaryotic serine/threonine or tyrosine protein kinases; in contrast, for all other phytochromes (conventional phytochromes) sequence similarities within the C-terminal domain to the catalytic domain of bacterial histidine kinases have been reported. We performed careful sequence comparisons of the putative catalytic domains of phytochrome with each other, with authentic serine/threonine, tyrosine and with histidine kinases. We report that conventional phytochromes exhibit structural elements of the catalytic domains of both histidine and, to a lesser extent, of serine/threonine and tyrosine kinases. The significance of these observations is discussed in the framework of the structure, function and evolution of phytochrome.**

*Key words:* Multiple sequence alignment; Phosphorylation; Plant signal transduction; Plant photoreceptor; Sequence conservation; Histidine kinase

## 1. Introduction

Phytochrome is a red/far-red light-regulated molecular switch which triggers a wide variety of photomorphogenetic responses in plants [1]. Phytochrome is encoded by a small multi-gene family; in *Arabidosis thaliana* five different members of this gene family have been identified [2]. Type I phytochrome, which is abundant in dark-grown or -adapted plants is a homodimeric protein. The apparent molecular weight of the monomer is about 120–127 kDa. Red light absorption of the phytochrome molecules is accomplished by an open chained tetrapyrrol chromophore. Per monomer, one chromophore molecule is covalently attached to the chromophore domain (highly conserved in all phytochromes) via a thio-ether bond. The initial step in phytochrome-dependent signal transduction chains is a reversible, red light-mediated isomerization of the chromophore [3]. The isomerization of the chromophore which interacts with the phytochrome apoprotein finally results in the formation of the biologically active form of phytochrome (Pfr).

The pathways of Pfr-dependent signal transduction are still not clear but there is growing evidence for the involvement of G proteins, $Ca^{2+}$ and cGMP in the regulation of certain Pfr-dependent reactions [4,5]. With phytochrome deletion mutants expressed in transgenic tobacco plants, Cherry et al. [6] showed that the C-terminal domain is required for phytochrome function. Based on sequence similarities to the catalytic domain of bacterial and eukaryotic protein kinases (PKs) Schneider-Poetsch et al. [7] and Thümmler et al. [8] proposed that the C-terminus of phytochrome has a catalytic activity. In the moss *Ceratodon purpureus* a gene has been identified which codes for a phytochrome (PhyCer), with a deduced mol. wt. of about 145 kDa, which has a C-terminus homologous to eukaryotic serine/ threonine and tyrosine protein kinases (EPKs). In vitro phosphorylation experiments with moss protonemata extracts revealed the presence of a 140 kDa protein which is phosphorylated in a red/far-red light-dependent manner [9]; and with expression experiments in fibroblast cell cultures of a chimeric protein composed of the extracellular domain of the epidermal growth factor receptor (EGF-R) and the PK catalytic domain of PhyCer, an auto-phosphorylation activity of PhyCer at serine and threonine residues could be demonstrated [10]. All other known phytochromes (which will here be referred to as 'conventional' phytochromes) do not exhibit homology to EPKs but rather to bacterial sensor proteins which are known to be histidine kinases (HKs) [11]. The existence of a HK catalytic domain as an integral domain of the phytochrome molecules is strengthened by the recent detection of HKs in yeast [12], animals [13] and in plants [14]. Sequence similarities between HKs and EPKs have not been identified. Accordingly, no sequence identities between the C-termini of conventional phytochromes and PhyCer were detected.

We performed a careful sequence comparison between the C-termini of different phytochromes and the catalytic domain of bacterial and eukaryotic PKs. Here we report the detection of sequence conservation between the C-terminal domains of conventional phytochromes and PhyCer and the catalytic domain of EPKs. These findings implicate the presence of elements of the catalytic domains of both the HKs, as well as the EPKs, within the C-terminus of conventional phytochromes. The significance of these observations is discussed in the framework of the structure, function and evolution of phytochrome.

## 2. Materials and methods

The multiple sequence alignment shown in Fig. 2 was made with the ALNED multiple sequence alignment editor. The ALNED program was written by David G. George [15]. Data base searches were done with the BLASTP program [16] at the NCBI using the BLAST network

*Corresponding author. Fax: (49) (89) 178 2274.


*SSDI* 0014-5793(94)01327-6

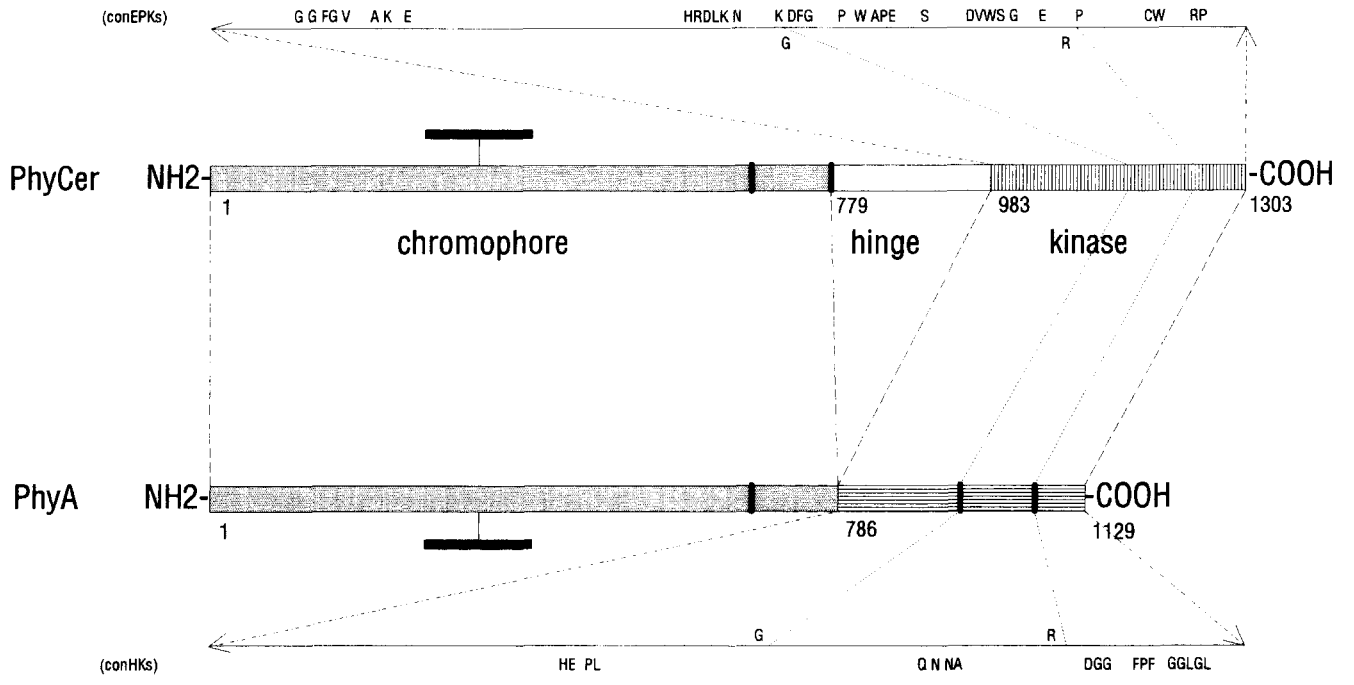



Fig. 1. Comparison of the structures of phytochrome from *C. purpureus* (PhyCer) and oat PhyA deduced from the respective gene sequences. The conserved chromophore domains are given as grey boxes (chromophore), the chromophore is symbolised by a black, horizontal bar. The box with vertical lines represents the EPK catalytic domain of PhyCer (kinase), which is linked to the chromophore domain by a hinge region (hinge; shown as a white box) with no sequence homology to any protein within the data bases. The box with horizontal lines represents the phytochrome domain with sequence similarities to the HK catalytic domain. Bold vertical lines mark the boundaries between the different exons of *phyCer* and *phyA* from oat. Numbers indicate amino acid residues of PhyCer and PhyA from oat. The EPK catalytic domain is enlarged on top of the structure of PhyCer, the HK catalytic domain is enlarged below the structure of PhyA. The highly conserved motifs of the catalytic domain of EPKs (conEPKs) and HKs (conHKs) are given above and below the figure. Conserved G and R residues which are found at the exon boundaries of conventional phytochrome sequences and which are also found at corresponding positions in PhyCer are indicated, too.

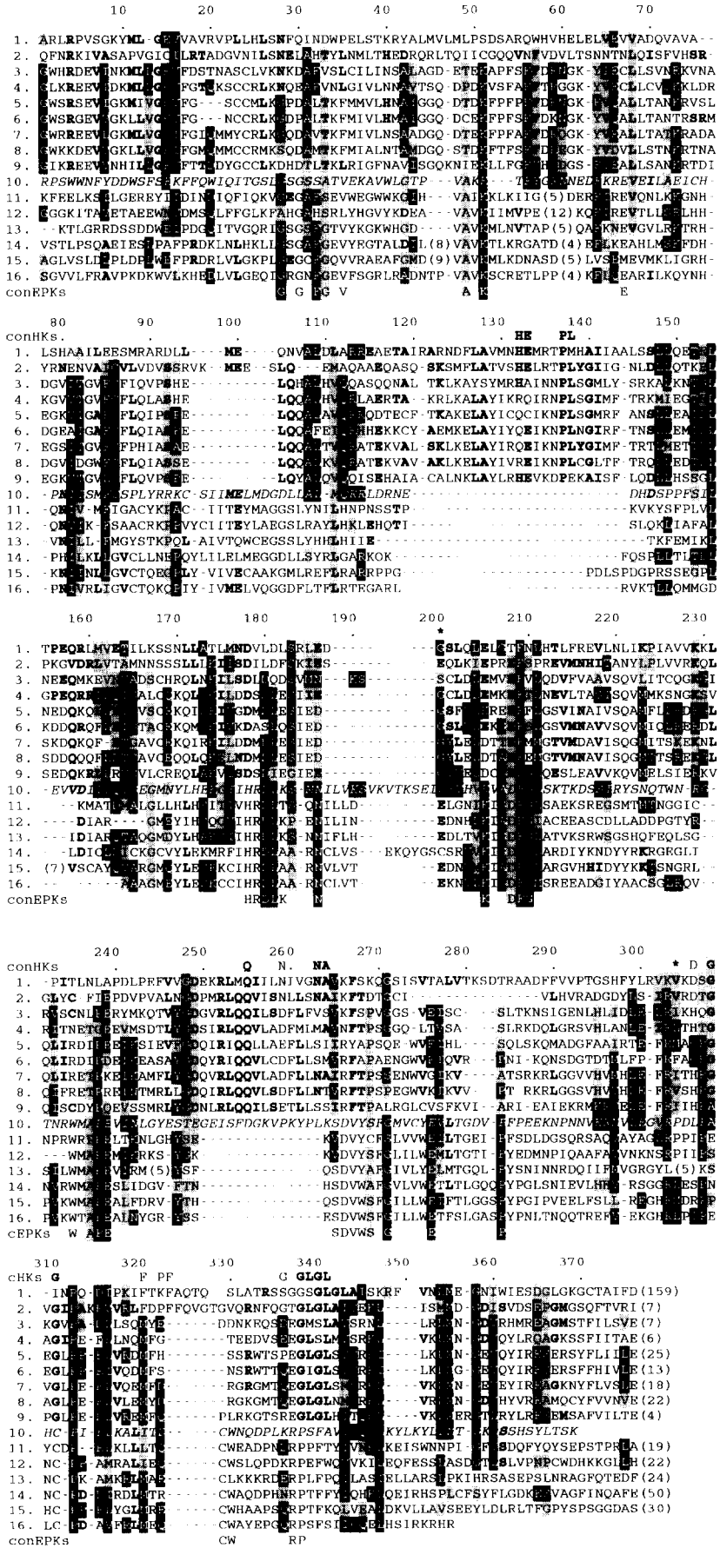
service. Sequence comparisons were performed with the FASTA [17] or the PILEUP program which is included within the GCG sequence analysis programs [18].

## 3. Results

In Fig. 1, the structures of conventional phytochrome (PhyA from oat) and PhyCer are compared. PhyCer is composed of three distinct domains; (i) the N-terminal chromophore domain of 778 amino acids (aa 1–778) which is highly conserved between all phytochromes, (ii) a domain of 204 aa (779–982) which exhibits no significant homology to any protein present within the protein databases (PIR, release 41; SWISS-PROT, release 29; and the cumulative weekly update to the major release; Brookhaven protein data bank, release April 1994; CDS translation from GenBank(R), release 84.0 and cumula-

→

Fig. 2. Multiple amino acid sequence alignment of the C-terminal domains of different phytochromes (3–10) and the catalytic domains of a bacterial (RCSC) (2) and a plant (ETR1) (1) HKs and eukaryotic Ser/Thr and Tyr-PKs (11–16). The sequence of PhyCer (10) is given in italics. The dashes represent gaps that have been inserted into the sequences to align conserved regions. Numbers and positions of amino acids not shown are given in parentheses. Identical amino acid residues present at corresponding positions in EPKs (including PhyCer) and in conventional phytochromes are given in black boxes. In a given column, these residues have to be present at least within two conventional phytochrome sequences as well as within two EPKs sequences (including PhyCer); structurally similar residues (see below), which are present within the same column and are present only in one phytochrome or in one EPKs sequences, are also given in black boxes (e.g. A in column 93). Identical residues in a given column, which are present only in one phytochrome and in one EPK sequence but are included in a motif of at least two residues, are also considered (e.g. ST; columns 218–219); additional identical residues in the same columns are given in black boxes as well (e.g. R in column 227). Structurally similar residues conserved in EPKs and conventional phytochromes are represented in grey boxes. They are taken into account only when identical residues are present within the same column, too. Similar groupings used for this purpose are non-polar chain R groups (M, I, L and V); aromatic or ring containing R groups (F and Y); small R groups with near neutral polarity (A, S, T, P and G); acidic and uncharged polar R groups (D, E, N and Q); and basic polar R groups (R, K and H). When there are no identical residues in a given column, conserved residues of the same amino acid group are considered when they are present at least within 10 sequences (excluding RCSC and ETR1) (e.g. T,P and A in column 49); when identical residues are also present, they are given in black boxes (e.g. A in column 42). Identical amino acid residues of RCSC or ETR1 present in at least two conventional phytochromes or two EPKs, are given in bold letters; identical motifs of at least two amino acids present only once in a HK and in phytochrome or in EPKs are also considered. Identical or similar residues, when given in black or grey boxes within the phytochrome and EPKs sequences, are given in black or grey boxes also within the HKs sequences. The highly conserved motifs characteristic for the Ser/Thr and Tyr-PKs (conEPKs or cEPKs) are represented below, highly conserved motifs characteristic for HKs (conHKs or cHKs) are represented above the alignment. The asterisks indicate the exon–intron boundaries between exon II and III (column 200) and III and IV (column 304) of the conventional phytochromes as described by Rüdiger and Thümmler [3]. Sequences and accession numbers under which the sequences can be found in the data bases are given in Table 1.



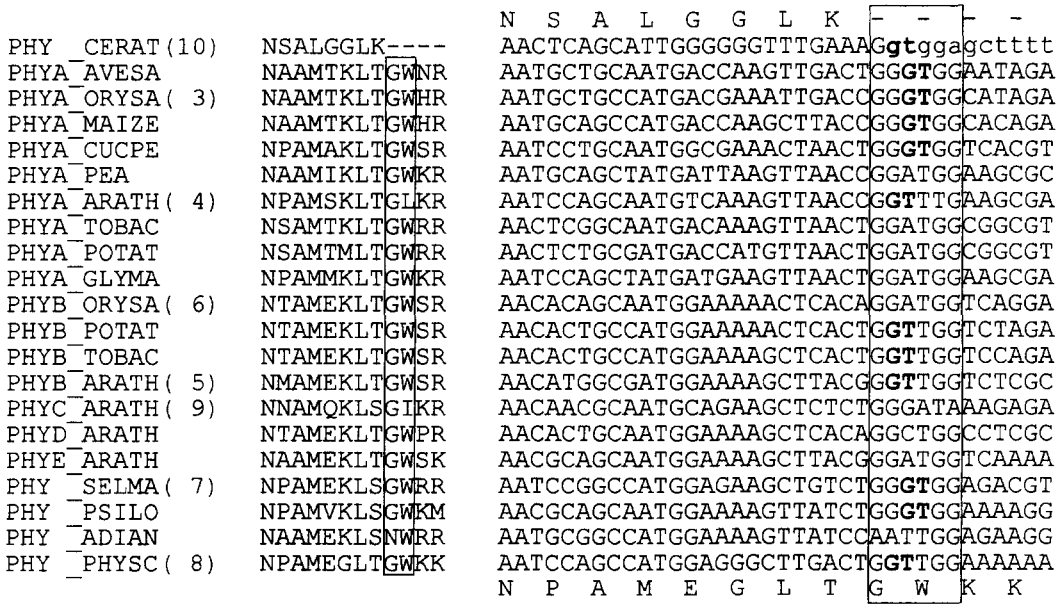



Fig. 3. Comparison of phytochrome sequences at the proposed transition point of the chromophore and the putative PK domains. In the left panel the different phytochrome types are indicated. The numbers in parentheses refer to the phytochrome sequences shown in Fig. 2. The middle panel shows the aligned amino acid sequences and the right panel the corresponding nucleotide sequences around the transition point of the chromophore and the putative PK domains. Only *phyCer* carries an intron at this position. On top and on the bottom of the nucleotide sequences the amino acid sequences of phytochrome of *Ceratodon* and of *Physcomitrella* are shown. Dashes within the *Ceratodon* sequence represent intron sequences. Boxed amino acid residues and nucleotide sequences indicate the highly conserved GW motif which marks the transition point of the chromophore and the putative PK domains. The nucleotide sequence GT, given in bold letters, represents potentional intron splice donor sites. EMBL accession numbers of phytochrome sequences not given in Table 1 are as follows: PHYA_AVESA, PhyA oat (m18822); PHYA_MAIZE, PhyA maize (PIR: JQ0382); PHYA_CUCPE, PhyA from *Cucurbita pepo* (m15265); PHYA_PEA, PhyA pea (x14077); PHYA_TOBAC, PhyA tobacco (x66784); PHYA_POTAT, PhyA potato (s84872); PHYA_GLYMA, PhyA soybean (l34842); PHYB_POTAT, PhyB potato (s51538); PHYB_TOBAC, PhyB tobacco (l10114); PHYD-ARATH, PhyD from *A. thaliana* (x76609); PHYE_ARATH, PhyE from *A. thaliana* (x76610; PHY_ADIAN, Phy from *Adiantum capillus* (d13519); PHY_PSILO, Phy from *Psilotum nudum* (x74930).

tive daily updates to the major release), and (iii) the C-terminal domain of 320 aa (983–1303) which is homologous to the catalytic domain of EPKs. The structure of conventional phytochromes, like PhyA from oat, differs from the structure of PhyCer. Conventional phytochromes are composed of only two domains; (i) the conserved N-terminal chromophore domain and (ii) a less-conserved C-terminal domain of about 340 aa with substantial sequence similarities to the catalytic domain of HKS (Fig. 1).

The comparison of C-terminal sequences of conventional phytochromes and PhyCer with different computer programs (BLASTP, FASTA, PILEUP) did not reveal any significant sequence similarities between conventional phytochromes and PhyCer. Nevertheless, by eye we could detect several amino acids which are conserved within the C-terminal 50 aa in all phytochromes, including PhyCer. Starting from this anchor and moving upstream to the N-terminus we aligned PhyCer with several phytochrome sequences by eye to maximise sequence identity. The alignment is given in Fig. 2. For comparison, the sequences of the catalytic domain of eukaryotic PKs closest to PhyCer are included. Also included in Fig. 2 are the sequences of a bacterial HK which is closest to conventional phytochromes (RCSC) [7,19] and of ETR1, the putative ethylene receptor in *A. thaliana* which exhibits substantial sequence similarities to bacterial HKs [14]. With this procedure we were able to detect amino acid motifs which are highly conserved at similar positions within the EPKs catalytic domain (including PhyCer) and the C-termini of conventional phytochromes (black-boxed residues in Fig. 2). The number of identical amino acids in Fig. 2 (see Table 1) is within the range of sequence conservation by chance between unrelated polypeptides; nevertheless, with the presence of stretches of conserved amino acid motifs at corresponding positions within the multiple sequence alignment, the likelihood that the sequence conservation is significant increases manifold [20]. Within the catalytic domain of EPKs, 35 aa have been identified which are highly conserved and which are necessary for phosphate transfer [21]. Of those, 14 residues are also found in conventional phytochromes; especially highly conserved are lysine (K) 51, the motif aspartic acid-leucine (DL) 180, phenylalanine (F) 209 and proline (P) 238. F209, which is included within the motif aspartic acid-phenylalanine-glycine (DFG), most indicative for EPKs, is present in all conventional phytochromes shown in Fig. 2 and in RCSC and in ETR1 as well; the aspartic acid of the DFG motif is substituted by the structurally similar residue glutamic acid (E) in the conventional phytochromes and in RCSC. The region corresponding to aa 175–215 of the alignment shown in Fig. 2 has been found to be conserved in other phosphotransferases which use ATP as a phosphate donor, like bacterial phosphotransferases which confer antibiotic resistance (e.g. hygromycin B phosphotransferase) [22]. The sequences have an invariant pair of aspartic acid residues resembling D180 and D208 in Fig. 2. Interestingly all phytochrome sequences more or less match to the consensus sequence described by Brenner [22], with the second aspartic acid residue being substituted by the similar glutamic acid residue (E208) in all conventional phytochromes. It is noteworthy that, within the alignment shown in Fig. 2, the region defined by Brenner exhibits the most

pronounced sequence conservation. This sequence conservation supports the proposal that phytochrome in general has phosphotransferase activity. Besides the aspartic acid within the DFG motif, of the 35 residues 4 are also substituted with similar residues in various phytochromes (G33, A49, A237 and S269). At positions which are not included within the consensus sequences of EPKs, residues can also be found which are highly conserved between EPKs and conventional phytochromes (L154, R303 and P312) or which are substituted with amino acids with similar structure (e.g. V and I at position 81; L,M,V and I at position 111; S,G,P and A at position 183 or I,V,M and L at position 343). The catalytic domain of HKs is less conserved; instead of 35 residues in EPKs, only 19 are highly conserved in HKs [11]. Of those 19 residues 13 are also found in conventional phytochromes, supporting the proposal that the catalytic domain of HKs and the C-terminus of conventional phytochromes are homologs [7]. None of the highly conserved residues in HKs is present in EPKs (including PhyCer). On the other hand several amino acid residues which are conserved between EPKs and conventional phytochromes are also present in RCSC and (or) in ETR1. In this respect, especially highly conserved residues are L154, F209, R303 and P312. Amino acids with similar structure which are highly conserved in HKs, EPKs and in phytochrome are, for example, L,M,V,I at position 111, S,G,P,A at position 183, E,N,D at position 186, L,I,V,M at position 204, I,V,L,M at position 316 and I,V,M at position 343. In most regions PhyCer exhibits higher sequence conservation with conventional phytochromes than authentic EPKs do. This is especially true around positions 56–60, 108–112, 200–203, 295–300 and 343–355 of the alignment, reflecting higher sequence conservation of PhyCer with conventional phytochromes than authentic EPKs with conventional phytochromes (see Table 1). Interestingly, however, in some regions the opposite is also true like, for example, at positions 1–20 or 247–248.

In Table 1 the number of identical and conserved amino acids between the catalytic domain of PKs and the C-termini of phytochrome are listed. From Table 1 (lanes A and B) it is evident that of the EPK catalytic domains, that of PhyCer is closest to conventional phytochromes (23% sequence identity of conventional phytochromes with PhyCer compared to 14–17% with authentic EPKs). Of those residues which are conserved between EPKs and conventional phytochromes (44–73 residues), only 29 and 27 are also found in ETR1 and in RCSC, respectively. Furthermore, conventional phytochromes are closer to EPKs (20–25% sequence identity) than HKs (ETR1 and RCSC) are to EPKs (11% and 13% sequence identity; see Table 1, lanes C and D). On the other hand, conventional phytochromes are closer to RCSC and to ETR1 than to EPKs (33% and 26% sequence identity of conventional phytochromes with RCSC and ETR1, respectively, compared to 20–25% with authentic EPKs). From all EPKs investigated, PhyCer is closest to HKs (7% and 9% sequence identity of ETR1 and RCSC, respectively, with PhyCer compared with 3–5% and 5–6% with authentic EPKs). It is interesting to note that the plant HK domain of ETR1 is less related to conventional phytochromes, which most likely also represent eukaryotic homologs of HKs, than the bacterial HK RCSC (26% sequence identity of ETR1 and 33% of RCSC to conventional phytochromes; Table 1, lanes C and D).

Table 1
Number of identical and similar residues conserved within the catalytical domains of EPKs (10–16), HKs (1 and 2) and the C-termini of various conventional phytochromes (3–9)

| Sequence | Species | A | B | C | D |
|---|---|---|---|---|---|
| 1. ETR1 | Arab. | 29 (9) | 44 (14) | 82[a] (26) | – |
| | | | | 34[b] (11) | – |
| 2. RCSC | *E. coli* | 27 (8) | 56 (18) | – | 107[a] (33) |
| | | | | – | 41[b] (13) |
| 3. PHYA | rice | 65 (20) | 91 (28) | 45[a] (14) | 57[a] (18) |
| 4. PHYA | Arab. | 63 (20) | 93 (29) | 54[a] (17) | 69[a] (22) |
| 5. PHYB | Arab. | 80 (25) | 101 (32) | 49[a] (15) | 59[a] (18) |
| 6. PHYB | rice | 73 (23) | 92 (29) | 55[a] (17) | 64[a] (20) |
| 7. PHY | Selag. | 64 (20) | 95 (30) | 55[a] (17) | 70[a] (22) |
| 8. PHY | Physc. | 70 (22) | 98 (31) | 51[a] (16) | 68[a] (21) |
| 9. PHYC | Arab. | 71 (22) | 90 (28) | 47[a] (15) | 60[a] (19) |
| 10. PHY | Cerat. | 73 (23) | 92 (29) | 21[b] (7) | 28[b] (9) |
| 11. DPYK2 | Dict. | 54 (17) | 72 (23) | 12[b] (4) | 19[b] (6) |
| 12. GmPk6 | soybean | 49 (15) | 61 (19) | 15[b] (5) | 19[b] (6) |
| 13. B-RAF | mouse | 44 (14) | 63 (20) | 16[b] (5) | 17[b] (5) |
| 14. C-ROS | chicken | 47 (15) | 69 (22) | 15[b] (5) | 18[b] (6) |
| 15. FGFR-4 | human | 49 (15) | 71 (22) | 13[b] (4) | 19[b] (6) |
| 16. FES | mouse | 50 (16) | 68 (21) | 11[b] (3) | 18[b] (6) |

A, number of identical residues (black boxes in Fig. 2) found at corresponding positions in EPKs (including PhyCer) and in conventional phytochromes; B, number of identical and similar residues (black and grey boxes in Fig. 2) found at corresponding positions in EPKs (including PhyCer) and in conventional phytochromes; C, number of identical residues found at corresponding positions in ETR1 and in [a]conventional phytochromes or [b]EPKs; D, number of identical residues found at corresponding positions in RCSC and in [a]conventional phytochromes or in [b]EPKs. Numbers in parenthesis represent the percentage of given residues with respect to the length of the sequence of the PK catalytic domain of PhyCer without gaps (320 aa) shown in Fig. 2. Arab., *Arabidopsis*; Selag, *Selaginella*; Physc., *Physcomitrella*; Cerat., *Ceratodon*; Dict., *Dictyostelium*. Sequences and accession numbers under which the sequences can be found in the EMBL/GenBank/DDBJ data bases are: (1) ETR1, HK, *Arabidosis* (124119); (2) RCSC, HK, *Escherichia coli* (m28242); (3) PHYA, rice (x14172); (4) PHYA, *Arabidopsis* (x17341); (5) PHYB, *Arabidopsis* (x17342); (6) PHYB, rice (x14065); (7) PHY, *Selaginella* (x61458); (8) PHY, *Physcomitrella* (x75025); (9) PHYC, *Arabidopsis* (x17343); (10) PHYC, *Ceratodon* (s51224); (11) DPYK1, Tyr-PK, *Dictyostelium* (m33784); (12) GmPk6, Ser/Thr-Pk, soybean (m67449); (13) B-RAF, proto-oncogene Ser/Thr-PK, mouse (m64429); (14) C-ROS, 'sevenless' homolog c-*ros*, Ser/Thr-PK, chicken (m13013); (15) FGFR-4, fibroblast growth factor receptor Tyr-PK, human (x57205); (16) FES, proto-oncogene Tyr-PK, mouse (x12616).

## 4. Discussion

The C-terminus of conventional phytochromes exhibits homology to the catalytic domain of bacterial HKs whereas in the case of PhyCer the C-terminus is homologous to the catalytic domain of EPKs. So far no sequence conservation between the catalytic domains of EPKs and of HKs has been observed. In Fig. 2 we show that the C-termini of conventional phytochromes and the C-terminus of *Ceratodon* phytochrome exhibit regions with highly conserved amino acid residues. Thus, the conventional phytochromes exhibit strong features of HKs and at the same time also features of EPKs. Because of the biliprotein nature of the chromophore domain (biliproteins normally serve as light-harvesting proteins in cyanobacteria) and the homology of the C-terminus of the conventional phytochromes to bacterial HKs, Thümmler [23] stated that phytochrome is of bacterial origin: this statement is based on the theory that the chloroplasts evolved from ancient endosymbiontic cyanobacteria (e.g. [24]). The evolution of the endosymbiontic bacteria to highly specific plant organelles was accompanied by a massive transfer of bacterial genes to the nucleus of the plant ancestor. It is now tempting to speculate that a former bacterial light-sensing apparatus captured features of eukaryotic signaling pathways within the eukaryotic environment. Since EPKs are essential components of most eukaryotic signalling pathways it is conceivable that a former bacterial HK adapted structural features of EPKs during evolution to form a PK hybrid; at the moment we do not know whether or not conventional phytochromes are capable of phosphorylating His, Ser, Thr, Tyr or any other amino acid residue, as we might expect from the gene sequences. In the case of PhyCer the situation is the opposite; the C-terminus of PhyCer is clearly homologous to the catalytic domain of EPKs. We can speculate that during evolution the former bacterial-type HK catalytic domain has been exchanged with an eukaryotic-type PK catalytic domain in PhyCer, probably by exon shuffling [25]. Following this line, the EPKs domain then adapted to the light-signalling cascade by acquiring some bacterial elements. Possibly, the regions specifically conserved between PhyCer and conventional phytochromes are involved in specific interactions with the conserved chromophore domain, thereby conferring light regulation to the activity of the PK catalytic domains. On the other hand, the sequence conservations shown in Fig. 2 could reflect only structural conservations resulting in similar peptide folding with no direct influence on the catalytic activity of the proteins. We can not answer this question at the moment. In a few other cases, the interaction of typical eukaryotic with typical prokaryotic signalling pathways was observed. SR-I rhodopsin, the light-sensing pigment in halobacteria, which is homologous to the animal-specific $\beta$-adrenergic receptor, is directly coupled to HtrI, which is homologous to eubacterial chemotaxis transducers like Tsr in *E. coli* [26]. HKs identified in plants and yeast are linked to cellular phosphorylating cascades via Ser/Thr EPKs. The putative ethylene receptor in *A. thaliana* (ETR1) acts upstream of CTR1, an EPK which is a negative regulator of the ethylene response [14]. For the HK SNL1, a yeast sensor for osmolarity, a linkage with the MAP kinase signalling pathway has been demonstrated [27]. Griffith et al. [28] reported sequence similarities between the phytochrome chromophore domains and BRLA, a regulator of asexual sporulation in the ascomycete fungus *Aspergillus nidulas*. In *A. nidulas*, asexual sporulation is regulated in a red/far-red light-dependent manner. As already mentioned, PhyCer was probably created by exon shuffling; in this respect we made an interesting observation: in Fig. 3, amino acid and nucleotide sequences within the region of the transition of the chromophore to the putative PK catalytic domains of all phytochrome sequences currently available are compared. Exactly at the 'fusion point', where *phyCer* carries an intron, all conventional phytochromes (with the exception of *phy* from the fern *Adiantum capillus*) exhibit a conserved glycine residue (G) followed by a highly conserved tryptophan (W; in 18 out of 20 phytochrome sequences). The codons for GW are GG(A/C/G/T) TGG; with the use of the codons GGT or GGG for the conserved G-residue, a splicing site donor motif, GT, is formed [29]. From 20 conventional phytochromes, 11 exhibit a potential splicing site at a corresponding position compared with *phyCer*. It is tempting to speculate that in conventional phytochromes quiescent splicing sites are present, allowing differential fusion of alternative catalytic domains to the conserved light-sensing domain of phytochrome. Another interesting observation can be made from Fig. 2; at the positions where introns are located in conventional phytochromes [3], highly conserved amino acid residues are found in conventional phytochromes, PhyCer and in RCSC or ETR1 (G at positions 200 and R at position 304 in Fig. 2; marked with asterisks; see also Fig. 1). We do not know if that observation has any relevance to phytochrome function or evolution.

Cherry et al. [6] reported that the terminal 35 aa residues in PhyA from oat are essential for physiological activity of overexpressed oat PhyA in tobacco. Interestingly, exactly within this region we find residues which are highly conserved in conventional phytochromes, PhyCer, HKs and also in some EPKs (at positions 343–360 in Fig. 2). Possibly, this short region has an important function for the activity of the different proteins. Because of the predicted amphipatic $\alpha$-helix nature of this region in conventional phytochromes, it was first speculated that it could be involved in phytochrome dimerization [30,31]; later, this hypothesis could not be confirmed with overexpression of deletion mutants of oat PhyA in transgenic tobacco [6] or by experiments with in vitro-translated fragments of oat PhyA performed by Edgerton and Jones [32].

Of course, we have to be careful interpreting data derived solely from sequence comparisons; especially the proposed PK catalytic activity of conventional phytochrome has now to be confirmed biochemically with purified phytochrome samples (this kind of experiment is currently in progress). Nevertheless, the presented data are surely helpful to the design of experiments which will be necessary to elucidate light-sensing signal pathways in plants.

*Acknowledgements:* We thank Dr. Andrzej Elzanowski (National Center for Biotechnology Information, Bethesda, MD 20894, USA) for helpful discussions. P.A. was supported by the Alexander von Humboldt Foundation, Germany. This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Bonn, Germany (Az Th 334/1-2).

## References

[1] Kendrick, R.E. and Kronenberg, G.H.M. (1994) Photomorphogenesis in Plants, 2nd Ed, Kluwer, Dordrecht.
[2] Clack, T., Mathews, S. and Sharrock, R.A. (1994) Plant Mol. Biol. 25, 413–427.

[3] Rüdiger, W. and Thümmler, F. (1991) Angew. Chem. Int. Ed. Engl. 30, 1216–1228.
[4] Neuhaus, G., Bowler, C., Kern, R. annd Chua, N.H. (1993) Cell 73, 937–952.
[5] Bowler, C., Neuhaus, G., Yamagata, H., Chua, N.H. (1994) Cell 77, 73–81.
[6] Cherry, J.R., Hondred, D., Walker, J.M., Keller, J.M., Hershey, H.P. and Vierstra, R.D. (1993) Plant Cell 5, 565–575.
[7] Schneider-Poetsch, H.A.W., Braun, B., Marx, S. and Schaumburg, A. (1991) FEBS Lett. 281, 245–249.
[8] Thümmler, F., Dufner, M., Kreisl, P. and Dittrich, P. (1992) Plant Mol. Biol. 20, 1003–1017.
[9] Algarra, P., Linder, S. and Thümmler, F. (1993) FEBS Lett. 315 69–73.
[10] Algarra, P., Herbst, R., Ullrich, A. and Thümmler, F. in preparation.
[11] Parkinson, J.S. and Kofoid, E.C. (1992) Annual Rev. of Gen. 26, 71–112.
[12] Ota, I.M. and Varshavky, A. (1993) Science 262, 566–569.
[13] Motojima, K. and Goto, S. (1993) FEBS Lett. 319, 75–79.
[14] Chang, C., Kwok, S.F., Bleecker, A.B. and Meyerowitz, E.M. (1993) Science 262, 539–544.
[15] George, D.G. (1987) PIR news letter No.3, National Biomedical Reserach Foundation, Washington, DC.
[16] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol 215, 403–410.
[17] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[18] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucl. Acids Res. 12, 387–395.
[19] Stout, V. and Gottesman, S. (1990) J. Bacteriol 172, 659–669.
[20] Taylor, W.R. (1988) Protein Eng. 2, 77–86.
[21] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42–52.
[22] Brenner, S. (1987) Nature 329, 21.
[23] Thümmler, F. (1992) In: Plant Photoreceptors and Photoreception: Proceedings of the Meeting of the British Photobiology Society, St. John's College Cambridge (Holmes, M.G. and Johnson, C.B. eds.) in press.
[24] Van den Eynde, H., De Baere, R., De Roeck, E.D., Van de Peer, Y., Vandenberghe, A., Willekens, P. and De Wachter, R. (1988) J. Mol. Evol. 27, 126–132.
[25] Gilbert, W. (1985) Science 228, 823–824.
[26] Spudich, J.L. (1993) J. Bacteriol. 175, 7755–7761.
[27] Maeda, T., Wurgler-Murphy, S.M. and Saito, H. (1994) Nature 369, 242–245.
[28] Griffith, G.W., Jenkins, G.I., Milnerwhite, E.J. and Clutterbuck, A.J. (1994) Photochem. Photobiol. 59, 252–256.
[29] Brown, J.W.S. (1986) Nucleic Acids Res. 14, 9549–9559.
[30] Partis, M.D. and Grimm, R. (1990) Z. Naturforsch. 45c, 987–998.
[31] Edgerton, M.D. and Jones, A.M. (1992) Plant Cell 4, 161–171.
[32] Edgerton, M.D. and Jones, A.M. (1993) Biochemistry 32, 8239–8245.