

Are Bacterial Sensory Systems Models for Phytochrome Action? Hydrophobic Cluster Analysis of the Phytochrome Module Related to Bacterial Transmitter Modules

Hansjörg A. W. Schneider-Poetsch, Christoph Sensen, and Sabine Hanelt
Botanisches Institut der Universität zu Köln, D-5000 Köln 41, Bundesrepublik Deutschland
Z. Naturforsch. **46c**, 750–758 (1991); received May 21, 1991

Phytochrome Action, Bacterial Sensors, Transmitter Modules, Hydrophobic Cluster Analysis, Structure Predictions

C-terminal amino acid sequences of two phytochromes sharing less than 40 per cent homology and two transmitter modules of bacterial sensor proteins that share homologies with phytochromes were examined by hydrophobic cluster analysis. Striking coincidences in distribution, size, shape and orientation of hydrophobic and hydrophilic domains of the bacterial transmitter modules and phytochrome sequences were revealed. The results corroborate the view that the folding of the homologous regions of the two groups of proteins is similar. Since the mode of action of phytochrome is not known, the structural coincidences may be indicative of how phytochrome transmits signals, although distinct differences should not be overlooked. Some coinciding structural features in phytochromes and bacterial sensors may tentatively be extracted from secondary structure predictions.

Introduction

Phytochrome sensing the environmental stimulus light is the only sensor of plants well known, isolated and sequenced. How this sensor works, however, is still a matter of discussion.

Recently we have reported on remarkable sequence homologies between the C-terminal region of phytochromes and a region of similar length (about 250 amino acids) of a group of bacterial proteins [1, 2]. The bacterial proteins sharing homologies with phytochromes belong to two-component systems composed of a sensor (possessing a transmitter module) and a regulator protein (possessing a receiver module). It is the transmitter regions of the sensor proteins that are homologous to the C-terminal regions of the light-sensors phytochrome. Bacterial sensors sense environmental stimuli like nitrogen limitation, or membrane composition by their N-terminal portions and transfer the signal by a conformational alteration of their transmitter module to a fitting regulator protein that in turn acts on enzyme activity or transcription. Light-mediated conformational alterations of the C-terminal parts of phytochrome have also been reported [3]. Hence, structural and functional coincidences infer that bacterial systems could

possibly be indicative of the mode of signal-transduction of phytochrome.

We have now extended our studies. Examining the homologous sequences by hydrophobic cluster analysis (HCA) [4] reveals that even regions with lower sequence homologies show a high degree of coincident hydrophilic and hydrophobic patterns. Up to 73 per cent of the hydrophobic amino acids are found in clusters similar in size, shape and orientation. The coincidences revealed by hydrophobic cluster analysis infer that the folding of these proteins is similar. They contribute to the view that the homologous domains of phytochromes and bacterial sensor proteins may satisfy comparable functions.

Materials and Methods

We compared bacterial transmitter modules with C-terminal sequences of phytochromes by means of hydrophobic cluster analysis (HCA) and secondary structure predictions.

The proteins compared by HCA [4] were the new C-terminal phytochrome sequences from *Se-laginella* [1] and *Zea* [2] and the transmitter modules of NtrB from *Bradyrhizobium* (swissprot: NTRB\$BRASP, P10578, sensing nitrogen limitation) [5] and of RcsC from *Escherichia coli* (swissprot: RCSC\$ECOLI, P14376, sensing capsule composition) [6].

Reprint requests to Prof. H. A. W. Schneider-Poetsch.

Verlag der Zeitschrift für Naturforschung, D-7400 Tübingen
0939-5075/91/0900-0750 \$ 01.30/0



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitalized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition "no derivative works"). This is to allow reuse in the area of future scientific usage.

The phytochrome sequences of the two species share sequence homologies of only 36 per cent. NtrBBR was selected because of the ample information available and RcsCEc because of its about 25 per cent sequence homology with both phytochrome sequences.

According to HCA of Gaboriaud *et al.* [4], a sequence is arranged (notwithstanding interruptions of the helix) α -helical. The helix is then cut parallel to the axis of the helix cylinder and unrolled. This representation is duplicated in order to make the sequence easier to follow and to give a better impression of the environment of each amino acid. Leucine, I, V, M, F, W, Y are considered as hydrophobic; A and C are considered as hydrophobic only in a hydrophobic environment. The sequences were arranged and flattened by "Net" provided by the software facilities "HUSAR" of EMBL Heidelberg. The printouts were duplicated and marked by hand.

Secondary structure predictions were performed by the algorithms of Garnier *et al.* [7], provided by PC/GENE (IntelliGenetics, 700 East El Camino Real Mountain View, California).

In addition to the proteins compared by HCA, the analysis was extended to the phytochromes of *Avena* [8], *Cucurbita* [9], *Pisum* [10], *Oryza* [11] and *Arabidopsis* [12], and the bacterial sensor proteins NtrBKp (sensing nitrogen limitation) [13], EnvZEc (sensing membrane composition) [14], CpxAec (sensing membrane composition) [15], PhoREc, PhoRBs (sensing phosphate limitation) [16, 17], low host range (LHR) and wide host range (WHR) VirAAt (sensing plant exudate) [18]. The bacterial proteins stem from *Bacillus subtilis* (Bs), *Escherichia coli* (Ec) and *Klebsiella pneumoniae* (Kp).

As to the present analyses, two other programs of PC/GENE, NOVOTNY and GGBSM, did not appear to supply results visibly superior to GARNIER.

Results

Hydrophobic cluster analysis

Comparing phytochromes and bacterial proteins bearing a transmitter module [19] on the basis of hydrophobic clusters [4] reveals surprising similarities between these two different groups of proteins. Taken in at a glance, hydrophobic and hy-

drophilic regions of the two analyzed phytochromes (*Zea*, *Selaginella*) and the two bacterial sensors (NtrBBR, RcsCEc) within the domains sharing sequence homologies are to a great extent coincident in position, size, shape and orientation (Fig. 1).

Excluding gaps, the homology scores (per cent overlapping hydrophobic clusters; see [4]) are about 82 per cent for the two analyzed phytochrome sequences and 67 per cent for the two bacterial proteins. The latter score is slightly surpassed by the figures comparing the two phytochromes and RcsC (70 and 73 per cent). The lowest scores (58 and 60 per cent) in this comparison, although significant, are shown by phytochromes and NtrB (see Table I).

Overlapping clusters between phytochromes and bacterial sequences decrease sharply in the regions preceding sequence homologies. The same is true for the pair NtrC and RcsC. Within these regions the homology scores indicate random coincidences or significant differences. Introducing gaps and arranging the sequences (particularly the long sequence of RcsC) according to coinciding clusters however, might reveal some more homologies even in regions where sequence homologies are not easily found. Two greater gaps (one for RcsC and one for phytochromes and NtrB) are demanded in the region of overlapping sequences.

Table I. Homology scores (2 · number of overlapping hydrophobic residues · 100%/sum of hydrophobic residues of the two compared sequences [4]) of phytochrome and bacterial transmitter modules. The analyzed proteins were NtrB of *Bradyrhizobium* spec., RcsC of *E. coli*, phytochromes of *Selaginella* and *Zea*. The scores are computed for the "module" and the region "preceding" the module region beginning with the first amino acid of NtrB (113 amino acids upstream from the module region; not fully shown on Fig. 1). (The relative frequency of hydrophobic amino acids in the analyzed sequences is about 37%; random sequences would score around $37 \pm 6\%$.)

Proteins compared	Module region	Preceding region
Sel-Zea	82.3	86.1
Sel-RcsC	73.1	40.5
Zea-RcsC	70.4	45.0
Zea-NtrB	59.6	31.2
Sel-NtrB	58.3	26.3
NtrB-RscC	67.4	20.8

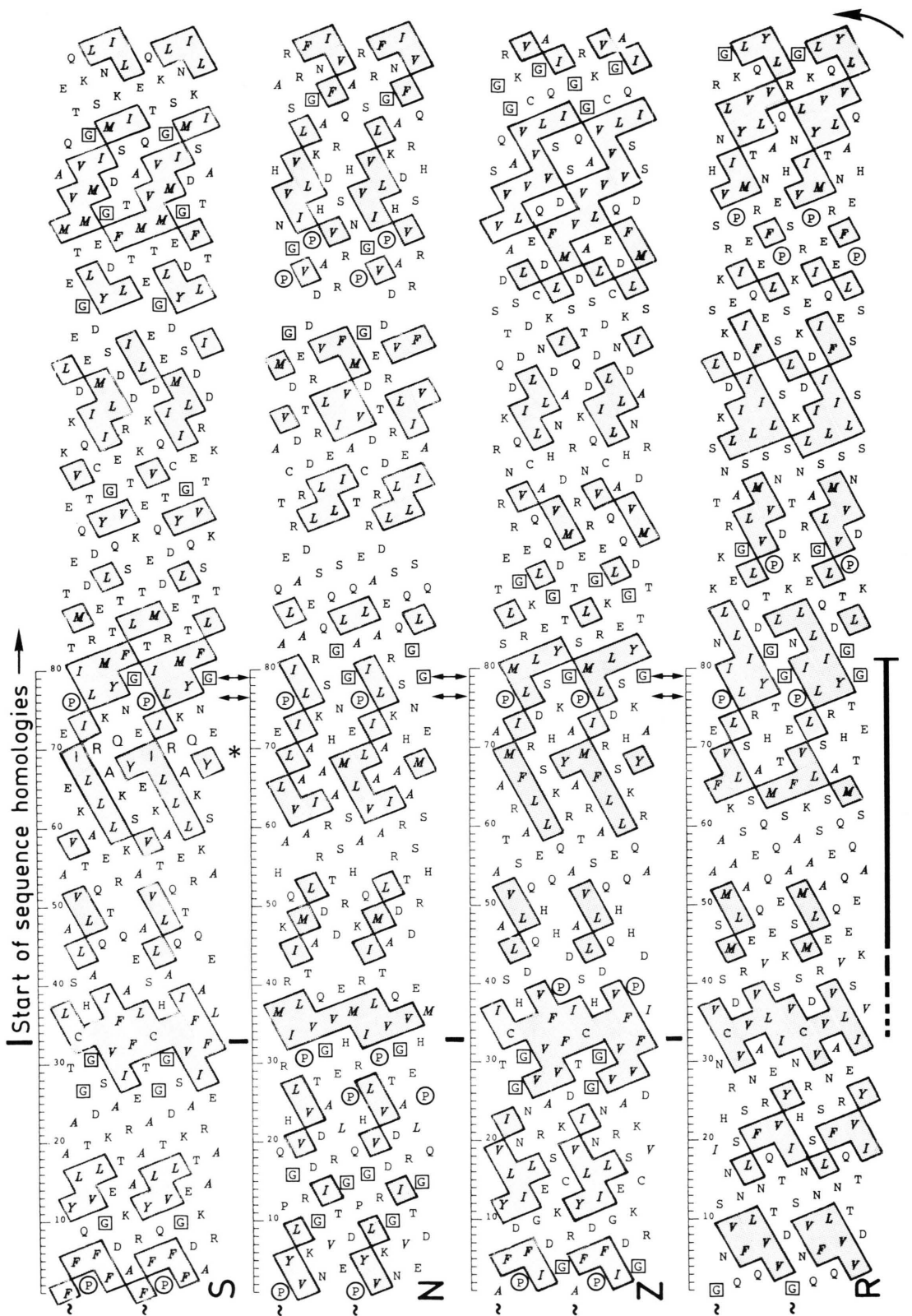
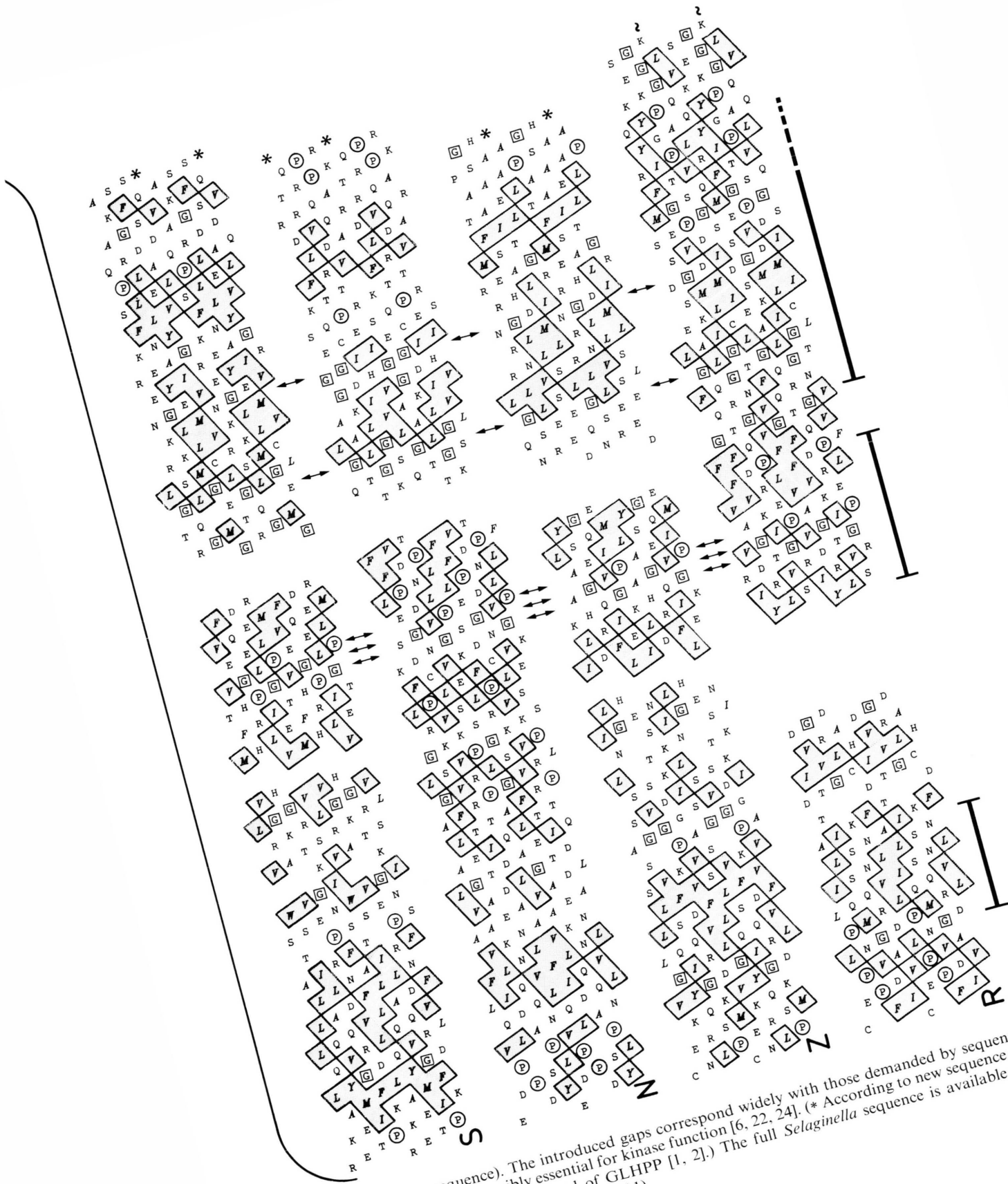


Fig. 1. Hydrophobic cluster analysis according to Gaboriaud *et al.* [4] of two phytochrome sequences (S = *Selaginella* [1], Z = *Zea* [2]), and two bacterial sensor proteins (N = NtrB from *Bradyrhizobium* [5], R = ResC from *E. coli* [6]). Note the coincidences in position, size, shape, and orientation of hydrophobic clusters (I, L, V, M, F, W, Y; shaded areas), and frequency and coincident positions (arrows) of glycines (boxed) and prolines (in circles); (~ = continuing



sequence. * = end of a sequence). The introduced gaps correspond widely with those demanded by sequence alignments [2]. Bars indicate areas possibly essential for kinase function [6, 22, 24]. (* According to new sequence data, the *Selaginella* sequence reads AYIRQ instead of GLHPP [1, 2].) The full *Selaginella* sequence is available from the EMBL data library (accession # X 61458; SMPHY 1).

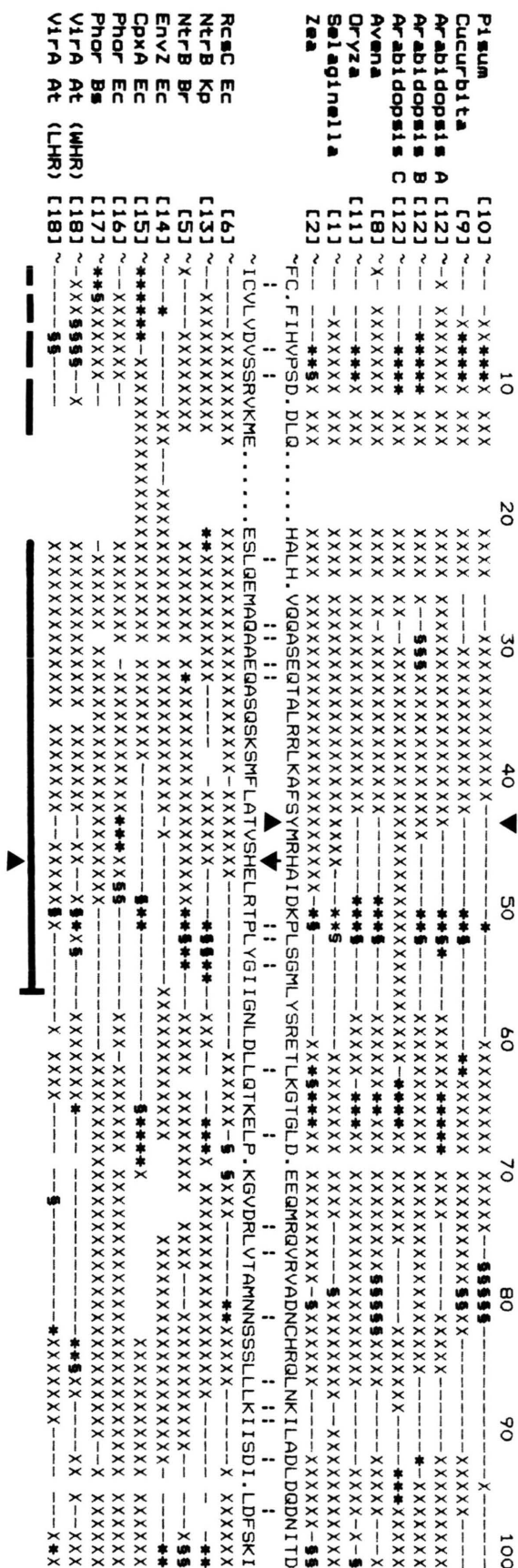
Glycines (and prolines) and a high percentage of hydrophilic residues in the non-interrupted sequence(s) point to loop regions. Loop regions may also be indicated in a few other places. In some regions the relative positions of prolines and glycines are highly conserved (arrows on Fig. 1). These latter parts of bacterial sensors coincide with regions assumed to be functionally important (vide infra).

Secondary structure predictions

To size up where the α -helix arrangement demanded by HCA is improbable, secondary structure predictions according to Garnier *et al.* [7] have been brought in. After analyzing the four proteins under investigation, a synoptical study appeared to be more pertinent: even the two phytochromes which can be assumed to have the same overall tertiary structures differed markedly in their predicted secondary structures. Therefore, seven more phytochromes and seven more bacterial sensors have been analyzed. Assessing these results, one has to remember that the accuracy of secondary structure predictions is assumed to reach at best 70 per cent [20] and that phytochromes are a more homogenous group of proteins than the listed bacterial sensors.

In the first third of the overlapping sequences, the helical arrangements prevail in both, the bacterial sensors and the phytochromes. In the middle section, helical and non-helical stretches are irregularly distributed, and in the final third, helical structures are preponderantly predicted for the phytochromes. Greater similarities in this section exist between *Selaginella* phytochrome, or the phytochromes of *Arabidopsis* and bacterial transmitters. On average, the phytochrome module shows 46.2 per cent helical and 36.6 per cent extended conformation, the bacterial module 40.0 (H) and 43.4 (E) per cent (NOVOTNY program).

As to coil and β -turn regions, some details may be of interest. Regions where such conformations are predicted for most of the analyzed proteins, are found around positions 50 (within mostly extended conformations) and 65, 150 and 160, 180 to 210, and 220 to 250 (see Fig. 2). The latter two regions include the two mentioned extended (vide supra) gaps in various sequences. Coils and turns are also found in the vicinity of gaps around position 70 to 80.



Discussion

According to the outlines and examples given by Gaboriaud *et al.* [4], the results allow to conclude that the compared phytochromes and bacterial sensor proteins are similar with respect to their overall folding in the regions matching the C-terminal 250 amino acids of phytochromes. This part of phytochrome fits the scheme of a transmitter module as defined by Kofoid and Parkinson [19] which is able to transmit signals to a receiver module.

Thus, by analogy a simple model describing the action of phytochrome can be deduced. In this model, P_{fr} (phytochrom transformed by red light, showing an absorption peak in the far red part of the spectrum) would conform to the active of the two interconvertible forms of the transmitter module protein (NtrB, P-2 on Fig. 2A) which influences the receiver module of its regulator protein (P-1 on Fig. 2A and 2B). To meet the more complex eukaryotic systems, one has to add structural features allowing the receiver module protein to travel through the membrane of the nucleus. Presently, there is no evidence that phytochrome itself has properties to enter this compartment.

Keener and Kustu [21] have pointed out that NtrB is a kinase although it does not have an easily recognizable ATP-binding site and sequence similarities to known protein kinases. It has been assumed that the kinase first auto-phosphorylates a histidine located within a sequence with homologies to AAHEIK [22] (see also [23] for chemotaxis proteins). There is also the notion that bacterial sensors contain the conserved sequence G.GLGLAI [6]. Sequences homologous to these sequences at coinciding positions are also contained in the C-terminal parts of phytochromes [2] (see centre of Fig. 2). Coil and turn structures are predicted for these regions. A histidine at the indicated position (amino acid 46 on Fig. 2), however, is not present in phytochromes of *Pisum*, *Cucurbita*, *Arabidopsis* (A and B), and *Selaginella*. In addition, the first G or rarely a S (amino acid 250 on Fig. 2) in the second motive is absent in all phytochromes. However, a tyrosine appears to be conserved in phytochromes only three residues upstream from the bacterial histidine. (Conserved amino acids liable to phosphorylation are additionally present in other regions and could be as-

sumed to be close to the module region in a dimeric molecule.)

In order to ascribe kinase function to the phytochrome module region, which certainly is an anchor serving a specific function (possibly in cooperation with other parts of this or another molecule), we are still in need of the structural details that constitute a kinase. Including the chemotaxis protein CheA of *Salmonella typhimurium* [25] in the analysis, where an alignment with the transmitters requires very large gaps, and assuming that all regions where sequence alignments demand gaps in sequences are of inferior significance, only a few regions seem to be essential. These are the regions spanning approximately the amino acids up to residue 56 (including the conserved Y, and other amino acids like R and C liable to phosphorylation), residues 160 to 177, and the amino acids downstream from residue 211 (numbering of Fig. 2; bars on Fig. 1 and 2).

Phytochrome-mediated phosphorylation would be in accord with current speculations (see [25]). By analogy and as to specificity, this process has to be assumed to be highly target-specific, although cross-talks between different sensory system of bacteria have been reported [26].

Analyses of the action of NtrB [21, 26, 27] point to the fact that sensors are parts of more complex systems. Kofoid and Parkinson [19] speak of dyadic relay systems where transmitters activated by a sensory input modulate receivers which themselves possess an output domain. An extract of essential details is given in Fig. 2B. NtrB (P-2) is a kinase/phosphatase [21] whose activity and specificity is allosterically regulated by a third protein (P-3, P_{II} of the literature) that interacts or does not interact with NtrB depending on the input signal modifying this (P-3) protein. It would be conceivable that phytochrome acts in cooperation with such a protein. However, analogies end by the fact that P_{II} is not modified by NtrB (the analogue to phytochrome) but by other proteins.

Although the models for phytochrome action are widely recommended by experimental data, unexpected coincidences and analogy, there are some more reasons to remain sceptical.

Firstly, peptides homologous to bacterial transmitter modules have in eukaryotic proteins so far only be detected in phytochromes [1, 2], where the "module" region is less conserved than other parts.

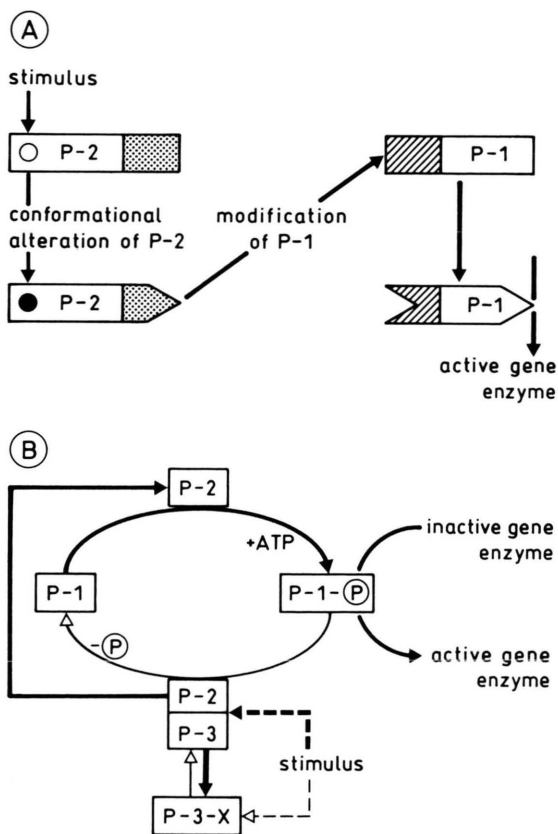


Fig. 3. Models for the action of phytochrome based on structural and functional homologies with bacterial sensor proteins and the bacterial transfer of information from an transmitter module to a receiver module. Both, bacterial sensors and phytochromes sense environmental stimuli by their N-terminal heads and confer signals by conformational alteration to their C-terminal parts (transmitter module) that share homologies. The signal can then be transferred to a fitting receiver module on a regulator protein.

A) Simple model: P-1 = regulator protein (receiver); P-2 = sensor protein (transmitter; NtrB, phytochrome); dark stippling = conserved domains of sensors (transmitter module); diagonal lines = conserved domains of regulators (receiver module).

B) Model in analogy to the action of NtrB [11, 13, 14]: P-1 is phosphorylated by P-2 and dephosphorylated when P-3 interacts with P-2; interaction of P-3 with P-2 depends on an input signal (by phytochrome?).

An explanation for a lower conservation could be a flexible response to the demands of mutating receiver modules: however, receiver molecules are

still not detected. (As to eukaryotic transmitter modules however, one may note that some eukaryotic proteins like the extracellular domain of the platelet-derived growth factor receptor [28] share sequence homologies of 10 to 14% with the full length of the phytochrome "module" region.)

Secondly, secondary structure predictions have added only restrictedly to the present analyses although a synoptical view can reveal some patterns shared by phytochrome and bacterial transmitter modules. The accuracy of secondary structure predictions is not very high and they do scarcely allow predictions on tertiary structures which would be necessary to assess in detail the role and the significance of particular structures in signal transduction. NMR techniques that have proved to be powerful tools in elucidating protein modules (see [30]) would certainly add to the value of the comparative hydrophobic cluster analyses presented here.

Nonetheless, the models can be hoped to promote forthcoming discussions and experiments. Probably, the path indicated by them is less surprising than the homologies and analogies found. The possibility of cross-talks between phytochromes and bacterial receivers should be probed.

The present hypotheses do not preclude the possibility that phytochrome may additionally interact with other regulative systems.

The finding that P_{fr} *e.g.* is liable to ubiquitination [31] could link phytochrome with a system that has been implicated in a variety of diverse functions like modification of structure and function of nuclear and ribosomal proteins and protein degradation (see [32]). Conjugation with ubiquitin appears to be reversible, resembling the reversible and regulative nature of phosphorylation. So, histones and phytochromes appear to be altered by similar mechanisms.

It might be that the phytochrome molecule conceals still more surprises than expected.

Acknowledgements

This study was made possible by a generous grant of the Volkswagen Stiftung to H. A. W. S.-P.

- [1] H. A. W. Schneider-Poetsch and B. Braun, *J. Plant Physiol.* **137**, 576–580 (1991).
- [2] H. A. W. Schneider-Poetsch, B. Braun, S. Marx, and A. Schaumburg, *FEBS Lett.* **281**, 245–249 (1991).
- [3] H. A. W. Schneider-Poetsch, B. Braun, and W. Rüdiger, *Planta* **177**, 511–514 (1989).
- [4] C. Gaboriaud, V. Bissery, T. Benchetrit, and J. P. Mornon, *FEBS Lett.* **224**, 149–155 (1987).
- [5] T. Nixon, C. W. Ronson, and F. M. Ausubel, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7850–7854 (1986).
- [6] V. Stout and S. Gottesman, *J. Bacteriol.* **172**, 659–669 (1990).
- [7] J. Garnier, D. J. Osguthorpe, and B. Robson, *J. Mol. Biol.* **120**, 79–120 (1978).
- [8] H. P. Hershey, R. F. Baker, K. B. Idler, D. L. Lissimore, and P. H. Quail, *Nucl. Ac. Res.* **13**, 8543–8559 (1985).
- [9] R. A. Sharrock, D. L. Lissimore, and P. H. Quail, *Gene* **47**, 287–295 (1986).
- [10] N. Sato, *Plant Mol. Biol.* **11**, 697–710 (1988).
- [11] S. A. Kay, B. Keith, K. Shinozaki, and N. H. Chua, *Nucl. Ac. Res.* **17**, 2865–2866 (1989).
- [12] R. A. Sharrock and P. H. Quail, *Genes Develop.* **3**, 1745–1757 (1989).
- [13] W. J. Buikema, W. W. Szeto, P. V. Lemley, W. H. Orme-Johnson, and F. M. Ausubel, *Nucl. Ac. Res.* **13**, 4539–4555 (1985).
- [14] D. E. Comeau, K. Ikenaka, K. Tsung, and M. Inouye, *J. Bacteriol.* **164**, 578–584 (1985).
- [15] R. Albin, R. Weber, P. M. Silverman, *J. Biol. Chem.* **261**, 4698–4705 (1986).
- [16] B. L. Wanner and B.-D. Chang, *J. Bacteriol.* **169**, 5569–5574 (1987).
- [17] T. Seki, H. Yoshikawa, H. Takahashi, and H. Saito, *J. Bacteriol.* **170**, 5935–5938 (1988).
- [18] B. Leroux, M. F. Yanofsky, S. C. Winans, J. E. Ward, S. F. Ziegler, and E. W. Nester, *EMBO J.* **6**, 849–856 (1987).
- [19] E. C. Kofoed and J. S. Parkinson, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4981–4985 (1988).
- [20] G. D. Fasman, in: *Proteins Form and Function* (R. A. Bradshaw and M. Purton, eds.), *Protein Conformational Predictions*, pp. 135–145, Cambridge 1990.
- [21] J. Keener and S. Kustu, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4976–4980 (1988).
- [22] V. Weiss and B. Magasanik, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8919–8923 (1988).
- [23] D. Wylie, A. Stock, C.-Y. Wong, and J. Stock, *Biochem. Biophys. Res. Commun.* **151**, 891–896 (1988).
- [24] A. Stock, T. Chen, D. Welsh, and J. Stock, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1403–1407 (1988).
- [25] B. R. Singh and P.-S. Song, *Photochem. Photobiol.* **52**, 249–254 (1990).
- [26] A. J. Ninfa and B. Magasanik, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5909–5913 (1986).
- [27] C. W. Ronson, B. T. Nixon, and F. M. Ausubel, *Cell* **49**, 579–581 (1987).
- [28] R. K. Gronwald, F. J. Grant, B. A. Haldeman, C. E. Hart, P. J. O'Hara, F. S. Hagen, R. Ross, D. F. Bowen-Pope, and M. J. Murray, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3435–3439 (1988).
- [29] A. C. R. da Silva and F. C. Reinach, *Trends Biochem. Sci.* **16**, 53–57 (1991).
- [30] M. Baron, D. G. Norman, and I. C. Campbell, *Trends Biochem. Sci.* **16**, 13–17 (1991).
- [31] J. Shanklin, M. Jabben, and R. D. Vierstra, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 359–363 (1987).
- [32] S. Jentsch, W. Seufert, T. Sommer, and H.-A. Reins, *Trends Biochem. Sci.* **15**, 195–198 (1990).