The Xanthopsins: a new family of eubacterial blue-light photoreceptors

Published in:
EMBO Journal

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
The xanthopsins: a new family of eubacterial blue-light photoreceptors


Department of Microbiology, E.C.Slater Institute, BioCentrum, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands and 2Department of Biochemistry, Physiology and Microbiology, Laboratory of Protein Structure and Function, State University of Ghent, Ledeganckstraat 35, 9000 Ghent, Belgium
1Present address: Department of Microbiology and Molecular Genetics, Health Science Center at Houston, University of Texas, 6431 Fannin, Houston, TX 77030, USA
3Corresponding author

Photoactive yellow protein (PYP) is a photoreceptor that has been isolated from three halophilic phototrophic purple bacteria. The PYP from Ectothiorhodospira halophilica BN9626 is the only member for which the sequence has been reported at the DNA level. Here we describe the cloning and sequencing of the genes encoding the PYPs from E.halophilica SL-1 (type strain) and Rhodospirillum salexigens. The latter protein contains, like the E.halophilica PYP, the chromophore trans p-coumaric acid, as we show here with high performance capillary zone electrophoresis. Additionally, we present evidence for the presence of a gene encoding a PYP homolog in Rhodobacter sphaeroides, the first genetically well-characterized bacterium in which this photoreceptor has been identified. An ORF downstream of the pyp gene from E.halophilica encodes an enzyme, which is proposed to be involved in the biosynthesis of the chromophore of PYP. The pyp gene from E.halophilica was used for heterologous overexpression in both Escherichia coli and R.sphaeroides, aimed at the development of a holoPYP overexpression system (an intact PYP, containing the p-coumaric acid chromophore and displaying the 446 nm absorbance band). In both organisms the protein could be detected immunologically, but its yellow color was not observed. Molecular genetic construction of a histidine-tagged version of PYP led to its 2500-fold overproduction in E.coli and simplified purification of the heterologously produced apoprotein. HoloPYP could be reconstituted by the addition of p-coumaric anhydride to the histidine-tagged apopPYP (PYP lacking its chromophore). We propose to call the family of photoactive yellow proteins the xanthopsins, in analogy with the rhodopsins.

Keywords: Ectothiorhodospira halophilica/photoactive yellow protein/Rhodobacter sphaeroides/Rhodospirillum salexigens/xanthopsins

Introduction

The photoactive yellow proteins (PYPs) constitute a new family of eubacterial photoreceptor proteins (Hoff et al., 1994b). Members have been isolated from the halophilic phototrophic purple eubacteria Ectothiorhodospira halophilica (Meyer, 1985), Rhodospirillum salexigens (Meyer et al., 1990) and Chromatium salexigens (Koh et al., 1996). PYP is the first eubacterial photoreceptor to be characterized in detail and has recently been shown to contain a unique chromophoric group: thiol ester linked p-coumaric acid (Baca et al., 1994; Hoff et al., 1994a). This is the first demonstration of a co-factor role for p-coumaric acid in eubacteria, previously only known from higher plants (Goodwin and Mercer, 1983). The pathway of biosynthesis of p-coumaric acid has been extensively studied in higher plants (Hahlbrock and Scheel, 1989), but no information is available on the conservation of this pathway in E.halophilica or other eubacteria. In higher plants, the two enzymes of central importance in the metabolic conversions relevant for p-coumaric acid are: phenylalanine ammonia lyase (PAL), which catalyses the reaction from either phenylalanine or tyrosine to p-coumaric acid, and p-coumaryl:CoA ligase (PCL), which activates p-coumaric acid through a covalent coupling to CoA, via a thiol ester bond (Hahlbrock and Scheel, 1989).

The PYP from E.halophilica is by far the best-studied member of this photoreceptor family. Its crystal structure has recently been re-determined at 1.4 Å resolution and shows that the protein has an α/β fold, resembling (eukaryotic) proteins involved in signal transduction (Borgstahl et al., 1995). Evidence has been obtained indicating that PYP functions as the photoreceptor for a new type of negative phototaxis response (Sprenger et al., 1993). Absorption of a blue photon (λmax = 446 nm) induces PYP to enter a cyclic chain of reactions (Meyer et al., 1987). This photocycle involves two intermediates and strongly resembles the photochemistry of the archaebacterial sensory rhodopsins (Meyer et al., 1987; Hoff et al., 1994c).

Recently, the ORF encoding PYP from E.halophilica BN9626 was cloned and sequenced (Baca et al., 1995). Here we report the cloning and the complete sequence of the pyp genes from E.halophilica SL-1 (the type strain) and Rs.salexigens, which is the first gene cloned from this organism, through reverse genetics. Directly downstream of the pyp gene in E.halophilica we located a gene encoding a CoA ligase homolog, suggesting a plant-like conversion of p-coumaric acid to its CoA derivative before linkage to PYP lacking its chromophore (apoPYP).

Previously, we have reported the presence of a single cross-reacting protein in a large number of eubacteria, with a highly specific polyclonal antibody against PYP (Hoff et al., 1994b). Here we report, using heterologous PCR techniques, the identification of a new PYP homolog in the genetically well-characterized Rhodobacter sphaeroides. This finding opens the way to molecular genetic
studies of the function of PYP. The *E. halophila* pyp gene was heterologously overexpressed in *Escherichia coli* and *R. sphaeroides*, yielding (mainly) apoPYP. The purification of a histidine affinity-tagged derivative of PYP from *E. halophila*, overproduced in *E. coli*, yielded a 2500-fold overproduction of apoPYP. Intact PYP, containing the p-coumaric acid chromophore and displaying the 446 nm absorbance band (HoloPYP) could be reconstituted by the addition of p-coumaric anhydride to the recombinant apoPYP as described for apoPYP (Imamoto *et al.*, 1995). These results will facilitate detailed biophysical studies on a protein with a unique set of characteristics: it is water soluble, photostable and its structure is known at 1.4 Å resolution.

---

**Fig. 1.** The *pyp* gene from *E. halophila* SL-1 with flanking regions. (A) Physical map of the chromosomal region containing the *pyp* gene. The cloned 2.4 kb *PstI* fragment, which is located on the 5.2 kb *EcoRI*- *SphI* fragment, is shown in detail, indicating the position of the *dada*, *pyp* and *pcl* genes. The open arrow indicates the direction of the genes. (B) DNA sequence of the 1.8 kb *PvuII*- *PstI* fragment containing a partial ORF1, the *E. halophila* *pyp* gene and a part of *EcoRI*. The derived amino acid sequences are given at the first position of each codon by the one letter code. The stop codon is indicated by an asterisk. The putative AT-rich promoter region (41 mol% GC) is underlined. Putative ribosome binding sites are doubly underlined and an inverted repeat is overlined. Underlined amino acids are part of a highly conserved motif in AMP-binding proteins (Fulda *et al.*, 1994). The bases indicated by a vertical arrow differ from the formerly published *E. halophila* BN9626 sequence (Baca *et al.*, 1994).
**Results**

*The pyp genes from E. halophila and R. salexigens*

The DNA sequence of a 1.8 kb *PvuII*-PstI fragment was determined (Figure 1A) and is shown in Figure 1B. The amino acid sequence of *E. halophila* PYP predicted on the basis of this sequence information is identical to the one determined by amino acid sequencing (Van Beeumen *et al.*, 1993), except for position 56 which is a Gln instead of a Glu, as also observed in the DNA sequence of the pyp gene from *E. halophila* BN9626 (Baca *et al.*, 1995). A potential AT-rich (41 mol% GC) promoter region can be identified upstream of the ORF encoding PYP (positions 60–103, Figure 1B), which may be essential for the formation of an open complex for initiation of transcription. Also, a potential ribosome binding site (RBS) is located directly upstream of the PYP ORF. Directly downstream of the PYP ORF an inverted repeat is located (positions 557–587, Figure 1B).

The pyp gene from *R. salexigens* is the first gene cloned from this bacterium. It was localized on a 1.4 kb *PvuI*-SalI chromosomal fragment. Sequence analysis of this fragment (Figure 2B) showed that it contains the entire ORF encoding PYP; the predicted amino acid sequence contains 125 amino acids and completely matches the amino acid sequence of this protein (Koh *et al.*, 1996). Upstream of the ORF, a potential AT-rich (35 mol% GC) promoter region (positions 638–680, Figure 2B) and ribosome binding site can be recognized, while directly downstream of the ORF an inverted repeat is present (positions 1134–1164, Figure 2B).

**Identification of a PYP homolog in R. sphaeroides**

Chromosomal DNA from *R. sphaeroides* 2.4.1. was used as template in a PCR with two primers homologous to conserved pyp sequences to yield a 0.3 kb product. The validity of the PCR product was confirmed by Southern hybridization experiments with *R. sphaeroides* chromosomal DNA under stringent conditions, using the PCR fragment as a probe. This revealed strong and specific hybridization signals (data not shown). The DNA sequence of the product showed that the encoding protein sequence was homologous to PYP from *E. halophila*, *R. salexigens* and *Chromatium salexigens* (Figure 3).

**Comparison of PYP sequences**

The complete amino acid sequences of the PYPs from *E. halophila*, *R. salexigens* and *C. salexigens* (Koh *et al.*, 1996) showed that they are highly conserved, with the exception of PYP from *R. sphaeroides*.
1996) are homologous, with 66% of the amino acids identical in all three sequences. This result enabled us to obtain the partial sequence of a PYP homolog from *R.sphaeroides* (see above). A partial alignment of these four sequences is shown in Figure 3. All proteins contain the Cys residue that in the *E.halophila* protein has been shown to bind covalently to the chromophore (Van Beeumen et al., 1993). From the 1.4 Å crystal structure of PYP it can be concluded that Tyr42, Glu46, Arg52 and to a lesser degree Thr50 and Tyr98, in the *E.halophila* PYP, are important for the protein–chromophore interactions that lead to the deprotonation of the *p*-coumaric acid molecule and result in the tuning of the absorbance of this cofactor to 446 nm (Baca et al., 1995; Borgstahl et al., 1995; Kim et al., 1995). These residues are all conserved in the PYPs from *E.halophila*, *R.salexigens* and *C.salexigens* (Figure 3), in line with the similarities between these proteins with respect to their absorbance spectrum and photochemical properties (Meyer, 1985; Meyer et al., 1990). In the sequence of the *R.sphaeroides* PYP homolog these six residues, of central importance for the binding of the chromophore, are also conserved, with the exception of Thr 50 (Figure 3). Furthermore, a strong conservation is observed in the sequence VIGKNFF, which forms a type II tight turn between the α4-helix and the β3-strand of PYP (Borgstahl et al., 1995).

**Analysis of pyp flanking regions**

The 1.8 and 1.4 kb chromosomal fragments from *E.halophila* and *R.salexigens* respectively, were examined for the presence of ORFs. In addition to the PYP ORFs presented above, this analysis indicates the presence of a large partial ORF (391 residues) downstream of the pyp gene from *E.halophila* (Figure 1B). This ORF was not found in the chromosomal fragment from *R.salexigens*. In line with this, comparison of the 1.8 and 1.4 kb chromosomal fragments from *E.halophila* and *R.salexigens* showed that the sequence similarity in these fragments is confined to the ORFs encoding PYP.

Upstream of the pyp gene from *E.halophila* SL-1 an ORF is located that shows significant homology to the *E.coli dada* gene, encoding the small subunit of the membrane bound iron–sulfur flavoenzyme D-amino acid dehydrogenase (Olsiewski et al., 1980), as was found in *E.halophila* BN9626 (Baca et al., 1994). The partial ORF downstream of the pyp gene from *E.halophila* was further analyzed by searching for sequence similarities with proteins in the SwissProt database. The most similar proteins were found to be a number of CoA ligases from various organisms with ~24% sequence identity and 48% similarity over a stretch of 400 amino acids (Table II). Furthermore, this putative pcl gene (see Figure 1A) shows, like the pyp gene, a high GC-bias in the wobble position of its codons, which is indicative of its functionality. In *R.salexigens* the ORF encoding a CoA liga homolog has not been found downstream from the pyp gene. This may suggest a larger intergenic region between pyp and the putative pcl in this latter organism. This is supported by a Southern blot, showing hybridization of *R.salexigens* chromosomal digests with the putative *E.halophila* pcl (M.K Phillips-Jones, unpublished observations).

**Identification of the chromophore of R.salexigens PYP**

The chromophore of *R.salexigens* PYP was identified as *p*-coumaric acid in the purified protein with high-
To overexpress PYP from *E. halophila*, a 0.45 kb *Ava*II fragment from pYAMA958 containing the *pyp* ORF, was inserted into the overexpression plasmid pT713 (Studier et al., 1990) to yield pTY13. After transformation of pTY13 to *E. coli* BL21, 50- to 100-fold overproduction of PYP was observed using Western blots and rocket immunoelectrophoresis (RIEP). However, absorbance spectra of the cytoplasmic fraction of these cells do not show an absorbance band at 446 nm, while this band was expected to be clearly visible on the basis of the concentration of PYP determined by RIEP (data not shown). This indicates that *E. coli* BL21/pTY13 mainly produces apoPYP, i.e. PYP without the chromophore.

In an attempt to obtain an overexpression system for holoPYP, the plasmid pAKT3 (see Table I), containing the same 0.45 kb insert with the *pyp* gene from *E. halophila*, was conjugated to *R. sphaeroides* DD13. Since this organism is phototrophic, like *E. halophila*, and therefore produces a large array of pigments, it may also synthesize *p*-coumaric acid. The DD13 strain is mutated with respect to synthesis of the photosynthetic apoproteins (Jones et al., 1992), reducing the absorbance of the associating pigments, thereby facilitating the observation of the expected absorbance band at 446 nm, caused by holoPYP. RIEP experiments showed that the transconjugant *R. sphaeroides* DD13/pART3 also produces PYP at levels 100-fold higher than *E. halophila* (data not shown). Approximately 50% of the PYP produced was associated with the membrane fraction from these cells. However, also in this case the expected absorbance band at 446 nm for holoPYP was lacking (data not shown).

A chimeric version of the *pyp* gene from *E. halophila* was cloned in *E. coli*, which allows one to isolate PYP by the presence of a histidine affinity tag in the gene product and to confirm the lack of the chromophore in PYP produced in *E. coli*. Surprisingly, *E. coli* M15/pHis (see Table I) overproduces PYP at levels of 50 mg/l culture per OD_{660} unit, as determined by RIEP (Figure 5A), which is ~2500-fold higher than *E. halophila* and ~50-fold higher than in the case of the two overexpression systems described above. Cell-free extracts from *E. coli* M15/pHis were used in Ni-affinity chromatography. This method yielded ~75% pure protein in a single step (Figure 5B). Incubation of the isolated histidine-tagged PYP with enterokinase yielded a product with a molecular weight indistinguishable from native *E. halophila* apoPYP (Figure 5B). The absorbance spectrum of the isolated histidine-tagged PYP shows that the typical absorbance band in the visible region of the spectrum is completely lacking (Figure 5C). This indicates that the protein produced in this *E. coli* strain is histidine-tagged apoPYP (HAP).

To demonstrate the usefulness of HAP for further biophysical studies on PYP, we reconstituted HAP with *p*-coumaric anhydride into holoPYP. The following observations showed that reconstitution of holoprotein was achieved: (i) spectral analysis showed an absorption band at 446 nm, which increased (to saturation) with a stepwise addition of the *p*-coumaric anhydride; (ii) analysis of absorbance spectra in time showed an increase at 446 nm and a decrease at 350 nm, in line with an increase of holoPYP concentration and a decrease of the anhydride concentration; (iii) purified reconstituted holoPYP showed an absorbance spectrum like that of purified native PYP (Figure 5C); (iv) reconstituted holoPYP can be reversibly bleached after absorption of light (data not shown). The
Table 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strains and plasmids</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli BL21</td>
<td></td>
<td>Studier and Moffat (1986)</td>
</tr>
<tr>
<td>E. coli TG1</td>
<td></td>
<td>Gibson (1984)</td>
</tr>
<tr>
<td>E. coli S17-1</td>
<td></td>
<td>Simon et al. (1983)</td>
</tr>
<tr>
<td>R. sphaeroides TG1</td>
<td></td>
<td>Raymond and Sistrom (1969)</td>
</tr>
<tr>
<td>R. sphaeroides DD13</td>
<td></td>
<td>Van Niel (1944)</td>
</tr>
<tr>
<td>Rs. salexigens WS 68</td>
<td></td>
<td>Jones et al. (1992)</td>
</tr>
<tr>
<td>E. halophila SL1</td>
<td></td>
<td>Drews (1981)</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pCHB500</td>
<td>pRK415 and pSH3 derivative, Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Benning and Sommerville (1992)</td>
</tr>
<tr>
<td>pART3</td>
<td>0.45 kb E. halophila AvaiI fragment cloned into pCHB500</td>
<td>this study</td>
</tr>
<tr>
<td>pQE30</td>
<td>RBSII, 6×His tag, ColEl ori, amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Qiagen</td>
</tr>
<tr>
<td>pHisp</td>
<td>0.42 kb E. halophila PCR product cloned into pQE30</td>
<td>this study</td>
</tr>
<tr>
<td>pT713</td>
<td>expression vector, T7 promoter, Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Gibco BRL</td>
</tr>
<tr>
<td>pTY13</td>
<td>0.45 kb E. halophila AvaiI fragment cloned into pT713</td>
<td>this study</td>
</tr>
<tr>
<td>M13mp18/19</td>
<td>M13mp1 derived phages, lacZ&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Messing and Vieira (1982)</td>
</tr>
<tr>
<td>pYAMA18</td>
<td>2.4 kb E. halophila PstI fragment cloned into M13mp18</td>
<td>this study</td>
</tr>
<tr>
<td>pYAMA958</td>
<td>1.8 kb E. halophila PstI fragment cloned into M13mp18</td>
<td>this study</td>
</tr>
<tr>
<td>pS16</td>
<td>1.4 kb Rs. salexigens PvuI–SalI fragment cloned into M13mp19</td>
<td>this study</td>
</tr>
</tbody>
</table>

masse of the histidine-tagged hol- and apoPYP were determined by ESMS to be respectively, 16.0081 and 15.8625 kDa. These values correspond well to the calculated molecular weights of 16.0081 and 15.8611.

**Discussion**

We report here the DNA sequence of two genes encoding proteins known to be yellow and photactive. The sequence of pyp from *E. halophila* SL1 (type strain) is identical to the sequence reported for the pyp gene from *E. halophila* BN9626 (Baca et al., 1994). In the flanking regions six differences between the two sequences were found, which in five cases did not lead to changes in amino acid residues (see Figure 1B); this indicates the close similarity but distinctness of these two strains. Interestingly, all silent mutations are from T in the *E. halophila* BN9626 strain to G or C in the *E. halophila* SL-1 strain. This may be explained by a slight difference in the overall GC-content between the two strains, which have been isolated from different environments; the BN9626 strain was isolated from the Wadi Natrun, Lake Abu Gabara near Bir Hooker, Egypt (Imhoff et al., 1978) and the type strain SL-1 from Summer Lake, OR, USA (Raymond and Sistrom, 1969). The GC-content of the cloned DNA fragments from *E. halophila* SL1 and *Rs. salexigens* was calculated to be 67.3 and 65.8% respectively, which matches well with the overall GC-content from these organisms, being 68.4% (Raymond and Sistrom, 1969) and 64 ± 2% (Drews, 1981) respectively. The lack of a signal peptide sequence upstream from the two pyp genes is in line with the intracellular localization of PYP in *E. halophila*, as determined with immuno-gold labeling experiments (Hoff et al., 1994b). Furthermore, the isoelectric points of the PYPs from *E. halophila* and *Rs. salexigens* are predicted to be 4.63 and 4.23 respectively. For *E. halophila* PYP, this parameter was experimentally determined to be 4.3 (McRee et al., 1986).

The sequence data for these two PYPs were used to design primers for the amplification of a fragment from chromosomal DNA by heterologous PCR, leading to the identification of a PYP homolog in *R. sphaeroides*. The PCR product obtained was used as a probe to clone the *R. sphaeroides* pyp gene. This gene encodes a protein of 124 residues, which cross-reacts with a polyclonal antiserum raised against *E. halophila* PYP (data not shown). The amino acid sequence of the *R. sphaeroides* PYP homolog is ~46% identical to the sequence of the PYPs from *E. halophila*, *Rs. salexigens* and *C. salexigens*, indicating that this PYP belongs to a different subgroup of the yellow proteins (R. Kort and S.M. Hoffer, unpublished observations). Since *R. sphaeroides* is genetically accessible, this opens up possibilities for genetic studies concerning the function of PYP. The identification of this PYP homolog raises the question whether the *R. sphaeroides* protein also binds a c-pcoumaric acid chromophore. The conservation of Cys69, Tyr42, Glu46, Arg52 and Tyr98 in the *R. sphaeroides* sequence suggests that this may indeed be so. This leads to the prediction that *R. sphaeroides*, in addition to its well-studied positive phototactic and chemotactic responses (for a review see Armitage, 1992), displays additional phototaxis response(s), based on PYP (see Sprenger et al., 1993). This prediction is currently being tested.

Directly downstream of the pyp gene from *E. halophila* an ORF is located that shows the highest sequence similarity to a range of CoA ligases (Table II), including *p*-coumaryl-CoA ligases. The putative *E. halophila* CoA ligase contains the motif TSGSTGTP (Figure 1B), which is conserved in all members of the AMP-binding protein family, of which the coumaryl-CoA ligases form a distinct subfamily (Fulda et al., 1994). This motif resembles the known loop-forming adenine-binding motif (Saraste et al., 1990). In plants, coumaryl-CoA ligase is of central importance in the metabolism of *p*-coumaric acid (Hahlbrock and Scheel, 1989). This suggests that in *E. halophila*, *p*-coumaric acid is likewise activated by the formation of a thiol ester bond with CoA. The importance of this
esterification was demonstrated by the fact that \textit{in vitro} reconstitution of holopYP was observed with the thio-phenyl ester of \textit{p}-coumaric acid and not with \textit{p}-coumaric acid (Imamoto \textit{et al.}, 1995). A further indication for a functional coupling of the \textit{pyp} and \textit{pcl} gene products is the presence of an inverted repeat between these two coding regions and the absence of a recognizable promotor sequence, directly upstream of the \textit{pcl} gene (see Figure 1B). This indicates that transcription of the \textit{pcl} gene occurs by readthrough of this inverted repeat from the promoter directly upstream of the \textit{pyp} gene.

The biosynthesis of \textit{p}-coumaric acid, which in plants can be performed in one step by phenylalanine ammonia lyase (Hahlbrock and Scheel, 1989), may consist of three consecutive steps in prokaryotes (compare the amino acid fermentation scheme of the anaerobic bacterium \textit{Clostridium sporogenes}; Bader \textit{et al.}, 1982). If so, an aromatic aminotransferase, a 2-keto-acid reductase and a dehydratase respectively, would be involved. In the first reaction, pyruvate may be the amino acceptor, as shown for many aminotransferases. The reformation of pyruvate would then be carried out by alanine dehydrogenase. Interestingly, the \textit{dada} gene upstream of the \textit{pyp} gene (Figure 1A), encodes an alanine dehydrogenase.

Based on the observations described above, one can conclude that the organization of the genes encoding the PYP sensory system is completely different from that of the only other well-studied class of bacterial photoreceptors: the archaeabacterial sensory rhodopsins. For sensory rhodopsin I (SR-I) it has recently been shown that transcription of the \textit{sopI} gene (encoding the SR-I apoprotein) is transcriptionally coupled to an ORF immediately upstream of the \textit{sopI} gene; this upstream ORF (the \textit{htrI} gene) encodes the signal transducer interacting with SR-I (Yao and Spudich, 1992; Ferrando-May \textit{et al.}, 1993; Spudich, 1994).

In the soluble protein fraction of \textit{Rs.salexigens} cells, we could detect the PYP chromophore \textit{p}-coumaric acid (8 pmol/mg soluble protein). This finding made a protocol available for straightforward screening of intact cells for the presence of this chromophore. This may be of great importance, since the nature of the chromophore in receptors for a large number of blue-light responses, observed in microorganisms as well as in plants, has not yet been elucidated (Senger, 1987). The amount of chromophore identified in \textit{Rs.salexigens} is equivalent to 0.1 \textmu g PYP per mg soluble cell protein, similar to the cellular content of PYP in \textit{E.halophila} (Meyer \textit{et al.}, 1985).

We propose to designate the family of PYPs 'xanthopins', which is derived from the Greek words \textita{ξανθός} (yellow) and \textita{πιπερις} (pepper).
Table II. Homology of the putative courmary-CoA ligase from E.halophila with CoA ligases from other organisms

<table>
<thead>
<tr>
<th>Enzyme (number of amino acids)</th>
<th>Organism</th>
<th>Identity (%)</th>
<th>Similarity (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoA ligase homolog (391)</td>
<td>E.halophila</td>
<td>100</td>
<td>100</td>
<td>this paper</td>
</tr>
<tr>
<td>Acetate-CoA ligase (660)</td>
<td>A.eutrophus</td>
<td>25.1</td>
<td>49.5</td>
<td>Prieffert and Steinbuechel (1992)</td>
</tr>
<tr>
<td>Acetate-CoA ligase (672)</td>
<td>M.soehngenii</td>
<td>20.6</td>
<td>47.5</td>
<td>Eggen et al. (1991)</td>
</tr>
<tr>
<td>Long-chain-fatty-acid-CoA ligase (558)</td>
<td>E.coli</td>
<td>26.3</td>
<td>51.5</td>
<td>Black et al. (1992)</td>
</tr>
<tr>
<td>Long-chain-fatty-acid-CoA ligase (700)</td>
<td>yeast</td>
<td>22.8</td>
<td>45.5</td>
<td>Duronio et al. (1992)</td>
</tr>
<tr>
<td>Courmary-CoA ligase (545)</td>
<td>potato</td>
<td>22.4</td>
<td>45.5</td>
<td>Becker-Andre et al. (1991)</td>
</tr>
<tr>
<td>Courmary-CoA ligase (563)</td>
<td>rice</td>
<td>25.3</td>
<td>49.2</td>
<td>Zhao et al. (1990)</td>
</tr>
</tbody>
</table>

Identity and similarity values are based on full length alignments made with the Genetics Computer Group package program BESTFIT using a gap weight of 3.0 and a length weight of 0.1.

Cloning of the E.halophila pyp gene

PCR-digested E.halophila chromosomal DNA was used as a template in a PCR-reaction with degenerated oligonucleotides YS-1 and YS-2 with the sequences AARAAYTTTYTTAARGA and GTCTATYTMATRCTCRAA respectively, as based on the PYP amino acid sequence (Van Beemen et al., 1993). PCR was performed with the enzyme Taq polymerase (HT Biotechnology, Cambridge, UK) for 30 cycles with 1 min denaturation at 94°C, 1 min annealing at 20°C and 1 min elongation at 70°C. Based on the sequence of the PCR product a new probe was constructed, completely homologous to the pyp gene in E.halophila. This probe was used to isolate a positive clone (pYAMA18) by screening a mini library of 2.4 kb PstI chromosomal fragments from E.halophila in phage M13mp18. A 950 bp PstI fragment from pYAMA18, containing the pyp ORF, was subcloned in M13mp18 to give pYAMA958.

Cloning of the R.salexigens pyp gene

The probe used to clone the pyp gene from E.halophila was used in heterologous Southern hybridization experiments with R.salexigens chromosomal digests. A mini library, containing sized PstI–SalI fragments in phage M13 was screened by hybridization with the same probe, leading to the identification of two positive clones. A 1.4 kb fragment containing the pyp gene was made blunt by Klenow treatment and reinserted into the Smal linearized phage M13mp19, yielding pSt16.

Sequencing

Both strands of the 1.8 kb E.halophila PstI–PstI fragment and the R.salexigens 1.4 kb PstI–SalI fragment were sequenced using universal and gene-specific oligonucleotides; the sequence strategies are indicated in Figures 1A and 2A. Sequence information was obtained by the dyeoxy chain termination method (Sanger et al., 1977), using [35S]dATP and a modified T7 DNA polymerase sequencing kit (Sequenase; US Biochemical Corporation, Cleveland, OH), as well as through the use of fluorescently labeled dyeoxy nucleotides and a thermostable Taq polymerase with the Dyideoxy terminator cycle sequencing kit (Applied Biosystems, Foster City).

Identification of the R.sphaeroides pyp gene

Chromosomal DNA from R.sphaeroides 2.4.1 was used as template in a PCR using 10 cycles of annealing for 1 min at 25°C and 25 cycles at 35°C. Denaturation and elongation were performed in all 35 cycles for 1 min at 95°C and 72°C respectively. Primers were based on known pyp sequences and restriction sites BamHI and HindIII (underlined) were introduced to enable directional cloning: GCCGATCCGGCTTCCGC- GCCATCCAGCTGAG (NTPYP1) and GCCGAACTTTCTGACCG- GCTTGGCAGAAGCC (CTTPYP1). The PCR product obtained was isolated from agarose gel and inserted into phages M13mp18/19. Both strands of the PCR product were sequenced. Hybridization of the PCR product with R.sphaeroides chromosomal DNA was performed as described (Engler-Blum et al., 1993).

Identification of the chromophore of Rs.salexigens PYP

A colorless Rs.salexigens culture, grown aerobically in the dark in Hutner modified medium as described (Hoff et al., 1994b), was diluted twice in the same medium and incubated anaerobically at 42°C in a completely filled 500 ml screw-cap bottle under illumination with 60 W
The soluble cell fraction of 500 ml of aerobically and anaerobically grown cultures was prepared as described (Hoff et al., 1994b). Proteins were precipitated with 10% (v/v) trichloro-acetic acid and washed once with demineralized water. Pellets were resuspended in 5 ml demineralized water and incubated overnight at pH 12 (leading to a complete solubilization of the proteins) to hydrolyze thiol ester bonds, followed by acidification to pH 2 with hydrochloric acid and acetic acid to neutralize the chromatography optimal extraction. Before extraction, protein concentrations were determined with the Bio-Rad protein assay, as described by the manufacturer. Chromatography extractions were performed by mixing thoroughly with 15 ml ethyl acetate, followed by 5 min of centrifugation at 120 g. The organic phase was washed twice with 5 ml demineralized water and dried by air. To substantiate the result of our analysis, the same chromatography experiment was carried out using the purified Rs.salexensis PYP (Meyer et al., 1990). Air-dried samples were dissolved in distilled water and injected in a 50 μm fused silica capillary TSP050375 (Compoitive Metal Services LTD) with an injection time of 0.2 min and injection pressure of 40 mbar. The sample was analyzed in 60 mV Tris/30 mM valeric acid pH 8.2, through a capillary with an effective length of 55 cm, at 25 kV and ~12 μA. On-column detection was performed at 284 nm (determined as the wavelength at which trans p-cumaric acid maximally absorbs in the Tris/valeric acid buffer), with a UVIS 200 detector (Linear, Fremont). As a reference trans p-cumaric acid (Sigma, St Louis, MO) was used. To confirm this identification, p-cumaric acid was also subjected to electrophoresis in 25 mM borax buffer, pH 9.0 at 25 kV and ~35 μA. The amount of detected trans p-cumaric acid was calculated from the peak area using the software Caesar for Windows version 4.02, 1990, Prince Technologies. As a reference, 11.0 nl of trans p-cumaric acid (Sigma) was injected in the concentration range from 2.5 to 75 μM, showing a linear relation to the detected peak areas.

**Construction of overexpression plasmids and overproduction strains**

A 0.45 kb Avall fragment from pYAMA958, containing the pyr ORF from *E.elhalophilus*, was ligated into the Smal-linearized overexpression plasmid pT713 (Studier et al., 1990) to yield pT713, which was transformed to *E.coli* BL21. Overexpression in pT713 is based on the strong viral T7 promoter φ10. The gene coding for the viral RNA polymerase is located on the chromosome of *E.coli* BL21, downstream of an inducible lac promoter (Studier et al., 1990).

A conjunctive broad host range overexpression system was constructed by ligating the 0.45 kb Avall fragment, described above, into the PstI polynucleotide site of pHCB500. pHCB500 is a broad host range vector, containing two promoters directly upstream of the polynucleotide site: the *E.coli* T77 promoter and the Pφ17 promoter that supports anaerobic expression of the lac gene from the cytosolic enzyme (Bennig and Sommer, 1992). The resulting plasmid pART3 was transformed into the conjunctive strain *E.coli* S17 and then transferred to *R.sphaeroides* DD1 (Jones et al., 1992) by conjugation on LB agar plates for 4.5 h. Transconjugants were selected on LB plates containing tetracycline (10 μg/ml), streptomycin (5 μg/ml) and kanamycin (20 μg/ml). The transconjugants were subsequently grown in liquid medium under semi-anaerobic conditions, allowing pigment synthesis.

A third overexpression system involved the heterologous overproduction of an affinity-tagged version of PYP from *E.elhalophilus* in *E.coli*. The expression vector was constructed by directional insertion of a PCR product into the expression plasmid pQE30 (Qagen, Hilden). The PCR product was obtained using pYAMA18 as template in a reaction with the following primers: GCCAGCGCCATCAGAATTGCAAGGTTGAG NTCCGCCGCCCTGGC (NTPY2), containing the BamHI site (underlined) and CTPYP1 (see above). Use of NTPY2 reflects in the presence of an enterokinase site in the recombinant protein, allowing proteolytic removal of the affinity tag. This tag is formed by six His residues, encoded by pQEE30 (Qagen). The PCR was performed using an annealing temperature of 64°C for 30 s and extension at 70°C for 30 s in 30 cycles. The resulting PCR product was digested with BamHI and HindIII, ligated into pQEE30 (Qagen) to yield pHesp and transformed to *E.coli* M15. The colonies, resistant against ampicillin (100 μg/ml) and kanamycin (25 μg/ml), were shown to contain the construct by colony PCR, using the two primers described above.

**SDS-PAGE, Western blotting and RIEP**

SDS–PAGE was performed in a Bio-Rad mini slab gel apparatus (Bio-Rad, Hercules, CA) according to Laemml (1970) as modified by Schägger and Jagow (1987) for improvement of resolution at 5–20 kDa range. Gels were stained with Coomassie brilliant blue G250. Western blotting and immunodetection were performed as described previously (Towbin et al., 1979; Hoff et al., 1994b). RIEP was carried out as described (Hoff et al., 1994b).

**Heterologous expression of PYP**

*E.coli* BL21pTY13 and *E.coli* M15pHisp were induced to express the heterologous gene by the addition of 1 mM IPTG to well-aerated cultures of exponentially growing cells at an OD₆₀₀ of 1. Cells were grown at 37°C in well-shaken Erlenmeyers, or in a well-aerated 1 L fermentor (New Brunswick Scientific, New Brunswick). Production of PYP in *R.sphaeroides* was induced by growing the organism semi-anaerobically in two-thirds filled, slowly shaking Erlenmeyers, using Luria Bertani broth with appropriate antibiotics. The resulting *E.coli* and *R.sphaeroides* cells were sonified three times for 1 min while cooled on ice, and centrifuged at 200 000 g for 3 h at 4°C to obtain a clear supernatant containing the overexpressed protein. Absorbance spectra of these fractions were measured with an Amino DW2000 spectrophotometer (SLM Instruments). In addition, these fractions were used for SDS–PAGE, Western blotting and RIEP analysis, as described above.

**Isolation and cleavage by enterokinase of histidine-tagged PYP**

Ultrafiltration supernatants from *E.coli* M15pHisp, induced with IPTG, were incubated with Ni-NTA resin for 1 h at 4°C, as described by the manufacturer (Qiagen). The resin was packed in a column and eluted, either by an imidazole gradient or by a pH gradient, as described by the manufacturer. The protein elution pattern was analyzed by measuring the absorbance of the eluting fractions at 280 nm. Cleavage of histidine-tagged apoPYP was performed at 37°C for 5–24 h using an enterokinase:PYP ratio of 1:50 (w/w).

**Reconstitution of holoPYP**

Reconstitution of the heterologously produced apoPYP was achieved by addition of the p-cumaric anhydride, dissolved in dimethyl formamide (DMF), as described for the reconstitution of the apoPYP, obtained from *E.elhalophilus* (Imamoto et al., 1995). The p-cumaric anhydride was synthesized as described (Imamoto et al., 1995).

**Mass spectrometry**

The integrity of histidine-tagged apoPYP and reconstituted histidine-tagged holoPYP was verified by electrospray mass spectrometry (ESMS). Typically, 20 pmol of protein was dissolved in 10 ml CH₃CN:water:formic acid (1:0.9:0.1: v/v) and injected into the electrospray source of a VG Bio-Q mass spectrometer (VG Organic, Altrincham, UK) at a flow rate of 6 ml/min, delivered by a Harvard Syringe Pump 11 (Harvard, South Natick, MA). Nine-second scans, covering the 650–1550 amu range, were accumulated during 2.5 min. The spectra were collected and processed using the masslynx software provided with the instrument.

**Acknowledgements**

The authors are very grateful to R.Kok and J.van Thor for their advice on the use of molecular genetic techniques and to J.van Dijk and W.Spijker for cloning the Rs.salexensis pyr gene. We thank X.Xu and H.Vonk for performing the capillary electrophoresis experiments and H.P.M.Fennema for synthesis of the p-cumaric anhydride. We would like to thank B.Poolman for his help with the initial PCR experiments. We are very thankful to T.E.Meyer for supplying information concerning oligonucleotides YS-1 and YS-2 and the purified Rs.salexensis PYP. W.D.Hoff was supported by the Netherlands Organization for Scientific Research (NWO) via the Foundation for Biological Research (BIOn). J.Van Beeumen is indebted to the National Fund for Joint Basic Research for financial support (Contract 32001891).

**References**

Note added in proof

Recent results cast doubt on our strain assignment in E. halophila.