

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/14588113>

# Sequence Evidence for Strong Conservation of the Photoactive Yellow Proteins from the Halophilic Phototrophic Bacteria Chromatium salexigens and Rhodospirillum rubrum †

ARTICLE in BIOCHEMISTRY · MARCH 1996

Impact Factor: 3.02 · DOI: 10.1021/bi951494t · Source: PubMed

---

CITATIONS

43

---

READS

7

7 AUTHORS, INCLUDING:



**Moonjoo Koh**

Chosun University

13 PUBLICATIONS 213 CITATIONS

SEE PROFILE



**Gonzalez Van Driessche**

Ghent University

45 PUBLICATIONS 973 CITATIONS

SEE PROFILE



**Wouter D Hoff**

Oklahoma State University - Stillwater

100 PUBLICATIONS 3,112 CITATIONS

SEE PROFILE

# Sequence Evidence for Strong Conservation of the Photoactive Yellow Proteins from the Halophilic Phototrophic Bacteria *Chromatium salexigens* and *Rhodospirillum salexigens*<sup>†</sup>

Monjoo Koh,<sup>‡,§</sup> Gonzalez Van Driessche,<sup>‡</sup> Bart Samyn,<sup>‡</sup> Wouter D. Hoff,<sup>||</sup> Terrance E. Meyer,<sup>⊥</sup> Michael A. Cusanovich,<sup>⊥</sup> and Jozef J. Van Beeumen<sup>\*,‡</sup>

Department of Biochemistry, Physiology and Microbiology, University of Gent, Ledeganckstraat 35, 9000 Gent, Belgium, Department of Biochemistry, University of Arizona, Tucson, Arizona 85721, and Department of Microbiology, E. C. Slater Institute, BioCentrum, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands

Received July 5, 1995; Revised Manuscript Received December 4, 1995<sup>®</sup>

**ABSTRACT:** The photoactive yellow proteins (PYP) have been found to date only in three species of halophilic purple phototrophic bacteria. They have photochemical activity remarkably similar to that of the bacterial rhodopsins. In contrast to rhodopsins, however, the PYPs are small water-soluble proteins. We now report the complete amino acid sequences of *Rhodospirillum salexigens* and *Chromatium salexigens* PYP which allow comparison with the known sequence and three-dimensional structure of the prototypic protein from *Ectothiorhodospira halophila*. Although isolated from three different families of bacteria, the PYP sequences are 70–76% identical. All three contain 125 amino acid residues, and no insertions or deletions are necessary for alignment. This is a remarkable result when it is considered that electron transfer proteins from these purple bacterial species are only 25–40% identical and that insertions and deletions are needed for their proper alignment. It thus appears that PYP has the same important function in each of the purple bacteria and that most of the amino acid residues are necessary to maintain structure and function. By most standards, PYP would be called a “slowly evolving protein”. *R. salexigens* PYP is uniquely degraded by proteolysis at low ionic strength, probably as a consequence of unfolding due to electrostatic repulsion of the excess negative charge. Therefore it may also be classified as a “halophilic protein”.

The photoactive yellow protein (PYP)<sup>1</sup> was first reported in *Ectothiorhodospira halophila* (Meyer, 1985) and has since been found in two other halophilic purple bacteria, *Rhodospirillum salexigens* (Meyer et al., 1991), and *Chromatium salexigens* (Meyer et al., unpublished results). Photoactivity was first reported by Meyer et al. (1987), and continuing kinetics analysis has shown that the photocycle is remarkably similar to the bacterial rhodopsins from *Halobacterium halobium* and related species (Meyer et al., 1989, 1991, 1993; Hoff et al., 1992; Miller et al., 1993; Ng et al., 1995; Van Brederode et al., 1995). PYP is partially bleached and red-shifted from 446 to 465 nm in less than 1 ns, and then it is completely bleached and blue-shifted to 356 nm in 100  $\mu$ s and finally returns to ground state in 260 ms. Although there are more intermediate states in the photocycles of rhodopsins, the kinetics of PYP are very similar to those of sensory rhodopsin in particular (Bogomolni & Spudich, 1991). PYP

is small (14 kDa) and water soluble, whereas bacteriorhodopsin is larger (25 kDa) and completely spans the membrane. The amino acid sequence of *E. halophila* PYP shows that the protein contains a single peptide chain of 125 amino acid residues and that the *p*-hydroxycinnamic acid chromophore is covalently bound to cysteine 69 via a thiol ester (Van Beeumen et al., 1993; Hoff et al., 1994a; Baca et al., 1994). The three-dimensional structure, refined at 1.4 Å resolution, indicates that PYP contains a central  $\beta$ -sheet flanked on either side by short loops and helices (Borgstahl et al., 1995), whereas an earlier report suggested the structure consisted of two orthogonal  $\beta$ -sheets (McRee et al., 1989). Thus, there is no structural relationship to bacteriorhodopsin and no similarity except in the photocycle. The functional role of PYP is unknown, but it has been suggested that it is involved in negative phototaxis in the presence of blue light (Miller et al., 1993), a role similar to that of *Halobacterium* sensory rhodopsin. In order to gain some insight into the properties of PYP and to learn what residues are conserved and might therefore be structurally and functionally important, we have determined the sequences of *R. salexigens* and *C. salexigens* PYP.

## EXPERIMENTAL PROCEDURES

**Protein Modification.** The chromophore of the yellow proteins from *C. salexigens* and *R. salexigens* was removed by reduction with DTT and subsequent carboxymethylation with iodoacetic acid as described by Crestfield et al. (1963). Desalting was performed by gel filtration through a Sephadex X75 column (3.2  $\times$  300 mm) on a SMART chromatographic

<sup>†</sup> This work was supported by grants from the National Institutes of Health (GM21277), the Belgian Fund for Joint Basic Research (Contract 2.0018.91) and the Flemish Concerted Research Action (12052293).

\* Address correspondence to this author at Vakgroep Biochemie, Fysiologie en Microbiologie, Ledeganckstraat 35, B-9000 Gent, Belgium. Telephone: 32-9-264 5109. Fax: 32-9-264 5338. E-mail: jozef.vanbeeumen@rug.ac.be.

<sup>‡</sup> University of Gent.

<sup>§</sup> Present address: Department of Chemistry, Chosun University, Kwangju, Seoul 501-759, Korea.

<sup>||</sup> University of Amsterdam.

<sup>⊥</sup> University of Arizona.

<sup>®</sup> Abstract published in *Advance ACS Abstracts*, February 1, 1996.

<sup>1</sup> Abbreviations: CS, *Chromatium salexigens*; EH, *Ectothiorhodospira halophila*; PYP, photoactive yellow protein; RS, *Rhodospirillum salexigens*.

system (Pharmacia, Uppsala, Sweden), equilibrated and eluted with 100 mM ammonium bicarbonate, pH 7.8.

**Electrophoresis and Blotting.** SDS-PAGE of native proteins was carried out on a 15% T/3% C gel, according to Laemmli (1970), in a Mighty Small SE250 electrophoresis unit (Hoefer, San Francisco, CA). The separated proteins were then electroblotted onto ProBlot membrane (Applied Biosystems, Foster City, CA) in a Bio-Rad Transblot Cell, using 50 mM Tris-borate as blotting buffer. After blotting, the membrane was stained with 0.03% Coomassie blue in 30% methanol, 60% water, and 10% acetic acid and destained with 45% methanol, 45% water, and 10% acetic acid.

**Enzymatic Digestions of the Yellow Protein from *Chromatium salexigens*.** Digestion of 2.6 nmol of carboxymethylated apoprotein in 100 mM ammonium bicarbonate, pH 7.8, was performed with Glu-C endoproteinase (Boehringer, Mannheim, Germany) for 5 h at 37 °C, using an enzyme to substrate ratio of 1/20 (w/w). Two nanomoles of carboxymethylated apoprotein was digested with Lys-C endoproteinase (Wako, Osaka, Japan) at an enzyme to substrate ratio of 1/17 in 100 mM ammonium bicarbonate, pH 7.8. Incubation at 37 °C lasted for 4 h. Asp-N endoproteinase (Boehringer, Mannheim, Germany) digestion was carried out on 1 nmol of carboxymethylated apoprotein in 100 mM ammonium bicarbonate, pH 7.8, at 37 °C for 4 h, using an enzyme to substrate ratio of 1/24.

**Enzymatic Digestions of the Yellow Protein from *Rhodospirillum salexigens*.** Asp-N endoproteinase (Sigma, St. Louis, MO) digestion was performed on 8.5 nmol of native protein in 50 mM Tris-HCl, pH 8.0, at an enzyme to substrate ratio of 1/60 for 5 h at 37 °C. Four nanomoles of carboxymethylated apoprotein in 200 mM Tris-HCl, pH 8.5, was cleaved with Glu-C endoproteinase (Boehringer, Mannheim, FRG) at an enzyme to substrate ratio of 1/28 at 37 °C for 4 h. Digestion with Lys-C endoproteinase (Wako, Osaka, Japan) was carried out on 3 nmol of carboxymethylated apoprotein in 50 mM Tris-HCl, pH 8.5, for 4 h at 37 °C, using an enzyme to substrate ratio of 1/20. Reduced protein (5 nmol) was digested with trypsin (Boehringer, Mannheim, FRG) in 100 mM ammonium bicarbonate, pH 7.8, and 2 M urea, using an enzyme to substrate ratio of 1/24, and incubated for 18 h at 37 °C.

**Protein and Peptide Purification.** Peptides obtained after Asp-N endoproteinase digestion on native protein from *R. salexigens* were separated on a PTC-C18 column (2.1 × 220 mm, Brownlee), using a 140A Separation system equipped with a 1000S photodiode array detector (Applied Biosystems, Foster City, CA). The eluents pumped at a combined flow rate of 200  $\mu$ L/min were 0.1% TFA in MQ-water (solvent A) and 0.1% TFA in 70% ACN (solvent B).

The generated peptides from other digestions on reduced or/and carboxymethylated apoprotein from *C. salexigens* and *R. salexigens* were chromatographed on a C2/C18 column (2.1 × 100 mm) using a Pharmacia SMART chromatographic system. The peptides were eluted with a gradient of 0.05% TFA in 100% MQ-water (solvent A) to 0.04% TFA in 70% ACN (solvent B) at a flow rate of 100  $\mu$ L/min.

The yellow protein from *C. salexigens* was further separated from contaminating proteins (see Discussion) by reversed-phase HPLC on a C2/C18 2.1/10 column. The chromatographic equipment and elution conditions were the same as described above.

**Amino Acid Sequence Analysis.** N-terminal sequence analyses were performed on a 475A, 477A, or 476A pulsed liquid sequenator, equipped with an on-line PTH-amino acid analyzer (all from Applied Biosystems, Foster City, CA).

**C-Terminal Analysis.** A few nanomoles of carboxymethylated apoprotein from both origins was separately treated with a mixture of 0.1  $\mu$ g of carboxypeptidase A (Boehringer, Mannheim, Germany) and 0.5  $\mu$ g of carboxypeptidase B (Boehringer, Mannheim, Germany) in 100 mM ammonium bicarbonate buffer, pH 7.5. Analyses of released amino acids were carried out on a 420A Derivatizer with an on-line 130A Separation System (both from Applied Biosystems).

**Mass Analysis.** Plasma desorption mass spectra were collected on a BioIon 20K Biopolymer Mass Analyzer (BioIon, Uppsala, Sweden) during  $10^6$  counts of the  $^{252}\text{Cf}$  source at an accelerating potential of 15 kV. Samples were applied in small volumes of 0.1% TFA in Milli-Q water and spin-dried on nitrocellulose coated targets wetted with 5  $\mu$ L of 0.1% TFA in methanol.

Electrospray mass spectrometry was performed on a Bio-Q quadrupole mass spectrometer equipped with an electrospray ionisation source (Fisons Instruments, Altrincham, U.K.). Ten microliters of sample solution in 50% acetonitrile/1% acetic acid were injected manually in the 10  $\mu$ L loop of the Rheodyne injector and pumped to the source at a flow rate of 5  $\mu$ L/min. The solvent, 50% acetonitrile/1% acetic acid, was delivered by a 140A Solvent Delivery System (Applied Biosystems). Scans of 12 s over the mass range of 400–1600 amu were collected over 2 min. Calibration of the scans was performed with 50 pmol of horse heart myoglobin (Sigma).

Matrix assisted laser desorption mass spectrometry was carried out on a ToFSpec instrument using a Nitrogen-laser (337 nm) and a linear Time-of-Flight tube of 65 cm (Fisons Instruments, Altrincham, U.K.). Scans were accumulated over 77 laser shots, using  $\alpha$ -cyanohydroxycinnamic acid as matrix. External calibration was realized using both gramicidin S and bovine insulin (Sigma); their molecular masses are, respectively, 1141.7 and 5274.5 Da.

## RESULTS

Evidence for the sequences of photoactive yellow proteins from *C. salexigens* (CSPYP) and *R. salexigens* (RSPYP) is presented in Figures 1 and 2, respectively. Quantitative data of sequence analyses and mass data are given in Tables S.I–S.XIII of the Supporting Information.

**Sequence Determination of CSPYP.** Two nanomoles of native PYP were subjected to N-terminal sequence analysis and 39 out of 46 residues could be identified unambiguously. The same run also revealed the presence of a polypeptide which could be identified as a homolog of the CheY chemotactic factor of *Escherichia coli* (Van Beeumen et al., to be published) and an unidentified contaminant protein. The presence of each protein species could also be demonstrated by SDS-PAGE and subsequent sequence analysis of the electroblotted bands (Figure 3, lane A). N-terminal sequence analysis of band A clearly identified the missing residues at positions 20, 24, and 26 of CSPYP as two aspartic acids and leucine, respectively. Digestion of the carboxymethylated apoprotein with *Staphylococcus aureus* V8 protease resulted in the separation shown in Figure 4. Fraction

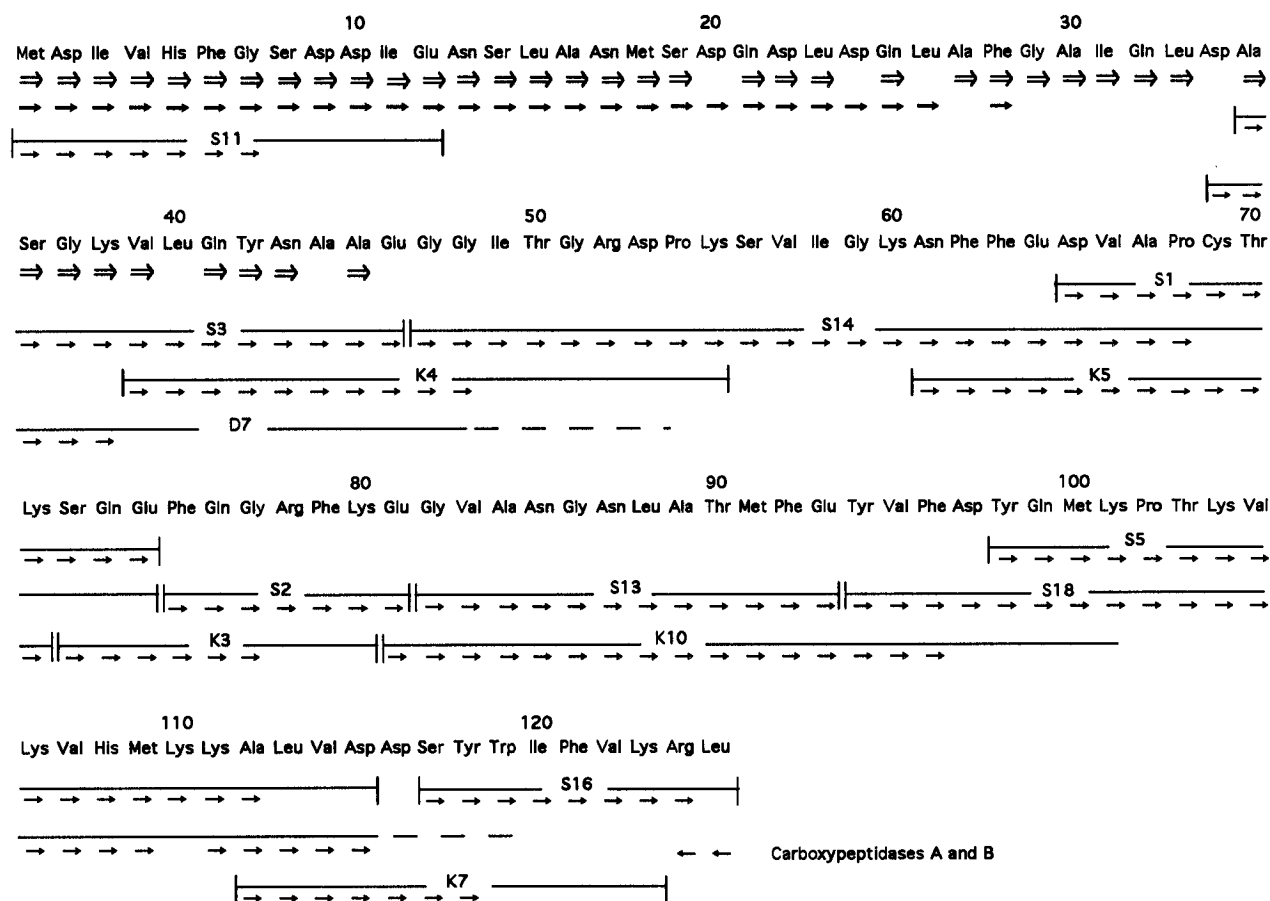
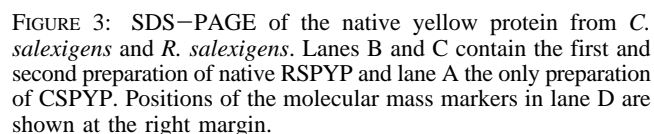
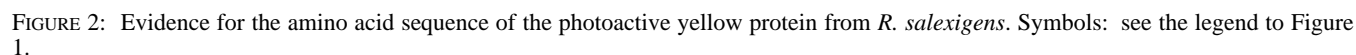


FIGURE 1: Evidence for the amino acid sequence of the photoactive yellow proteins from *C. salexigens*. Amino acids identified by Edman degradation of the native proteins are indicated by a double arrow, electroblotted proteins by arrows in bold, and peptides by simple arrows.

S14 was 28 residues long and contained one noncleaved Glu–Asp peptide bond. Partial cleavage of this bond resulted in fraction S1. The fifth residue of S1 was identified unambiguously as carboxymethylcysteine and extended fraction S14 by six residues. With fraction S3, we established the nature of the missing residues at positions 40, 44, and 46 as leucine, alanine, and glutamic acid, respectively. Peptide S5, which apparently originated by nonspecific cleavage of fraction S18, revealed the residue at position 110 as lysine. Digestion of the carboxymethylated apoprotein with Lys-C and Asp-N endoproteases resulted in the peptides shown in Figure S.1 and S.2. Connection of peptide S3 with S14 was established by sequence analysis of peptide K4. Fraction K3 was nine residues long and overlapped with peptides S1 and S2. Sequence analysis of peptide K10 showed a connection of S2 with S13 and an overlap with the first three residues of S18. K7 was sequenced for seven cycles with identification of an aspartic acid residue at position 116 and showed an overlap with the last five and first two residues of S18 and S15, respectively. Finally, the nature of the missing residue at position 34 could be determined as aspartic acid by N-terminal sequence analysis of fraction D7. Evidence for the C-terminal sequence was established by treatment of the carboxymethylated apoprotein with carboxypeptidase A and B as given by Figure S.3A.

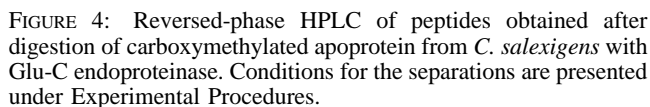
**Sequence Determination of RSPYP.** We used two samples of protein, one prepared with and one prepared without proteolytic inhibitors. Electrophoresis and blotting of the first protein sample obtained from Tucson revealed the presence of several polypeptide species (Figure 3, lane B), starting at different positions of the same protein species:

Met1 (band A), Ile4 and Phe6 (band B), Asp17 and Asp20 (band C), Ala21 and Ile23 (band D), and Asp 25 (band E). Those from the second protein preparation (Figure 3, lane C) started off at Gly8 (band A), Asp17 (band B), and Ala21 (band C). Due to strong proteolysis of the first protein mixture, prepared in the absence of inhibitors, we continued the sequence work on the second preparation. Identification of an aspartic acid at position 34 was used for digestion of the native protein with Asp-N endoprotease (Figure 5). Fraction D15 contained two peptide fragments of which one, D15A, was six residues long and appeared to be bound to the chromophore. The other, longer peptide, D15B, extended the already determined sequence from the blot (lane B) by 11 residues and identified the missing residue at position 38 as threonine. All other peptide fragments could be positioned by alignment with the sequence of CSPYP. In order to confirm overlaps of different Asp-N peptides, the carboxymethylated apoprotein was digested with Glu-C endoprotease and trypsin (Figures S.4 and S.5). The overlaps and/or connections were supported by sequence analysis of fractions S11, S7, S6, T8, T6, and T1. Although some overlaps in the regions 92–94 and 110–112 remained weakly proven, correctness of the sequence was completely supported by accurate mass analysis of all the peptides involved and by similarity with the sequences of the two other yellow proteins. Further, we also confirmed the presence of a cysteine residue at position 69 by sequence analysis of peptide K8, obtained after digestion of the carboxymethylated apoprotein with Lys-C enzyme (Figure S.6). Using carboxypeptidases A and B, the C-terminal



sequence was shown to be Phe Val Lys Arg Val (Figure S.3B).

We have determined the complete amino acid sequences of the new PYPs by N-terminal degradation of the electroblotted proteins and of enzymatically generated peptides. Both yellow proteins are 125 amino acids long. The measured mass (14 140.0 Da) (Figure 6A) of the yellow protein from *C. salexigens* fits very well the calculated mass



(13 993.9 Da), supplemented with the mass of the chromophore (146.2 Da). The nature of the chromophore was

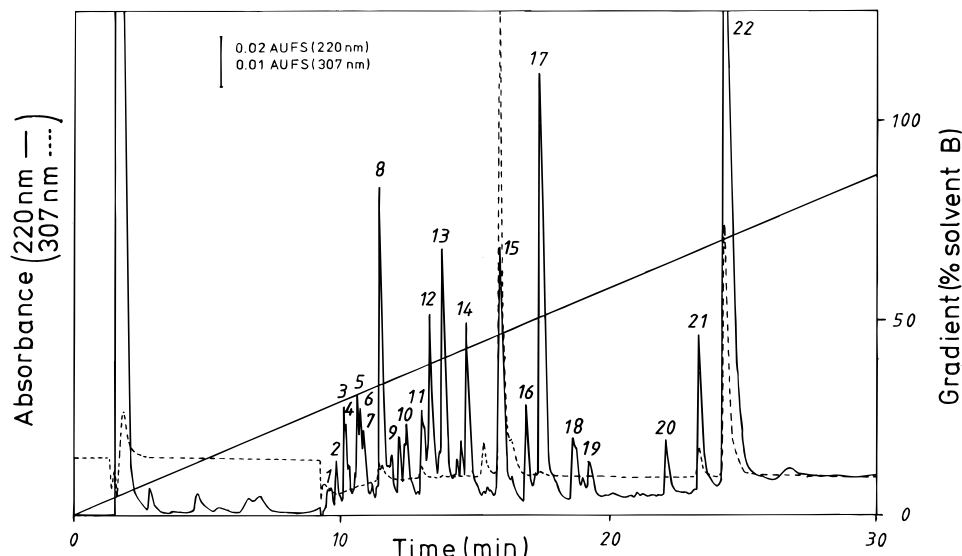


FIGURE 5: HPLC separation of peptides generated by digestion of native yellow protein from *R. salexigens* with Asp-N endoproteinase. Conditions are given under Experimental Procedures.

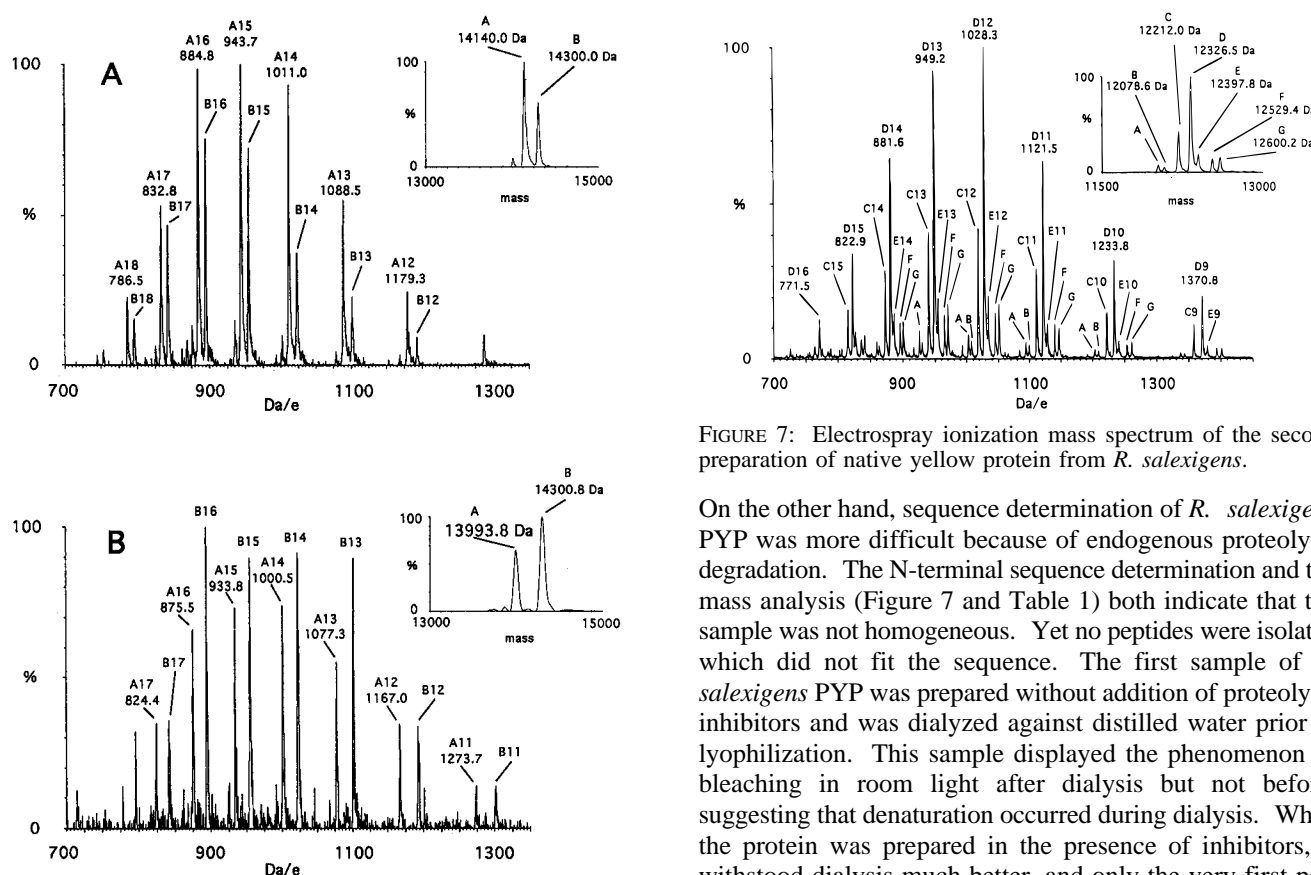


FIGURE 6: Electrospray ionization mass spectrum of native (A) and reduced (B) yellow protein from *C. salexigens*. The insert shows the transformed spectrum of the protein components as calculated by the MassLynx software, delivered with the instrument. The number following the capital letter at the top of each peak represents the number of positive charges for that particular  $m/z$  peak. The measured masses of 14 300.0 Da (A) and 14 300.8 Da (B) correspond to the theoretical mass of the contaminant chemotactic protein (Van Beeumen et al., manuscript in preparation).

determined by reduction of the native CSPYP with DTT and subsequent electrospray ionization mass measurement (Figure 6B), resulting in a mass decrease of 146.1 Da. This value was also obtained from a similar experiment on the native yellow protein from *E. halophila* (Van Beeumen et al., 1993).

FIGURE 7: Electrospray ionization mass spectrum of the second preparation of native yellow protein from *R. salexigens*.

On the other hand, sequence determination of *R. salexigens* PYP was more difficult because of endogenous proteolytic degradation. The N-terminal sequence determination and the mass analysis (Figure 7 and Table 1) both indicate that the sample was not homogeneous. Yet no peptides were isolated which did not fit the sequence. The first sample of *R. salexigens* PYP was prepared without addition of proteolytic inhibitors and was dialyzed against distilled water prior to lyophilization. This sample displayed the phenomenon of bleaching in room light after dialysis but not before, suggesting that denaturation occurred during dialysis. When the protein was prepared in the presence of inhibitors, it withstood dialysis much better, and only the very first part of the N-terminal region was ragged. We believe that the explanation for this phenomenon is that the protein is halophilic, that is, it requires high ionic strength for stability due to the large negative charge ( $-9$  from the sequence). Once unfolded, it is readily degraded by endogenous proteolytic enzymes. Therefore, we believe that native *R. salexigens* PYP has the same number of residues as does *E. halophila* PYP and *C. salexigens* PYP. We did not observe proteolysis with either *E. halophila* or *C. salexigens* PYP, which may be due to the smaller net charges ( $-7$  and  $-4$ , respectively) and therefore less electrostatic repulsion at low ionic strength. There may also be fewer proteolytic enzymes in *E. halophila* and *C. salexigens* which are sulfur bacteria

Table 1: Ragged N-Termini of Yellow Protein from *Rhodospirillum salexigens*, As Deduced from the Electrospray Ionization Mass Analyses

	measured (Da)	calculated with chromophore	difference (Da)	sequence position
First Preparation				
A	11 706.9	11 708.2	-1.3	23-125
B	11 906.8	11 907.4	-0.6	21-125
C	12 021.9	12 022.5	-0.6	20-125
D	12 078.5	12 079.6	-0.6	19-125
E	12 210.1	12 210.8	-0.7	18-125
F	12 324.5	12 325.9	-1.4	17-125
G	12 394.6	12 396.9	-2.3	16-125
H	12 527.4	12 528.2	-0.8	15-125
I	12 599.1	12 599.2	-0.1	14-125
Second Preparation (Figure 7)				
A	12 022.3	12 022.5	-0.2	20-125
B	12 078.6	12 079.6	-1.0	19-125
C	12 212.0	12 210.8	+1.2	18-125
D	12 326.5	12 325.9	+0.6	17-125
E	12 397.8	12 396.9	+0.9	16-125
F	12 529.4	12 528.2	+1.2	15-125
G	12 600.2	12 599.2	+1.0	14-125

as opposed to *R. salexigens* which utilizes a variety of organic compounds for growth (Drews, 1981). The photocycle of *E. halophila* PYP in distilled water is normal, suggesting no denaturation (Meyer et al., 1987; Hoff et al., 1994b).

The sequences of EHPYP, CSPYP, and RSPYP are compared in Figure 8. There is a very large 70-76% sequence identity without any insertions or deletions required. This is a remarkable fact when one considers that the three PYPs were isolated from different bacterial families. These bacterial species are known to have very few electron transfer proteins in common (Meyer, 1985; Meyer et al., 1990; Meyer et al., unpublished results). As an example of the sort of variation to be expected, cytochromes *c*<sub>2</sub> from species of Rhodospirillaceae average 40% identity and require numerous insertions and deletions for alignment (Ambler et al., 1979; Moore et al., 1990). One explanation, namely, that these species were not properly typed and are actually more closely related overall than their classification suggests, may be dismissed because electron transfer proteins isolated from these species are different from one another but similar to those of other species within each family (Ambler et al., 1994a, 1994b).

Another explanation for the extraordinary similarity of the PYPs is recent gene transfer, which in itself would not be too surprising since the three species live in similar habitats (Biebl et al., 1981; Pfenning et al., 1981; Trüper et al., 1981) and have similar GC contents of their DNA (Mandel et al., 1971). This possibility intrinsically assumes that the gene transfer would have occurred nearly simultaneously from one species to the other two since the sequence difference is roughly the same among all three yellow proteins. Thus, we consider this explanation to be highly improbable as well.

A third possible explanation for the remarkable similarity in species of PYP is that the structure is highly conserved because the proteins have similar essential functions which would be impaired or lost if there were any greater sequence divergence. This explanation, which for this particular protein implies that some three quarters of the sequence is required to maintain structure and function, appears to be most likely and defines PYP as a "slowly evolving" protein.

The N-terminal half of the PYP sequence is slightly more variable than is the second half. This may be due to a

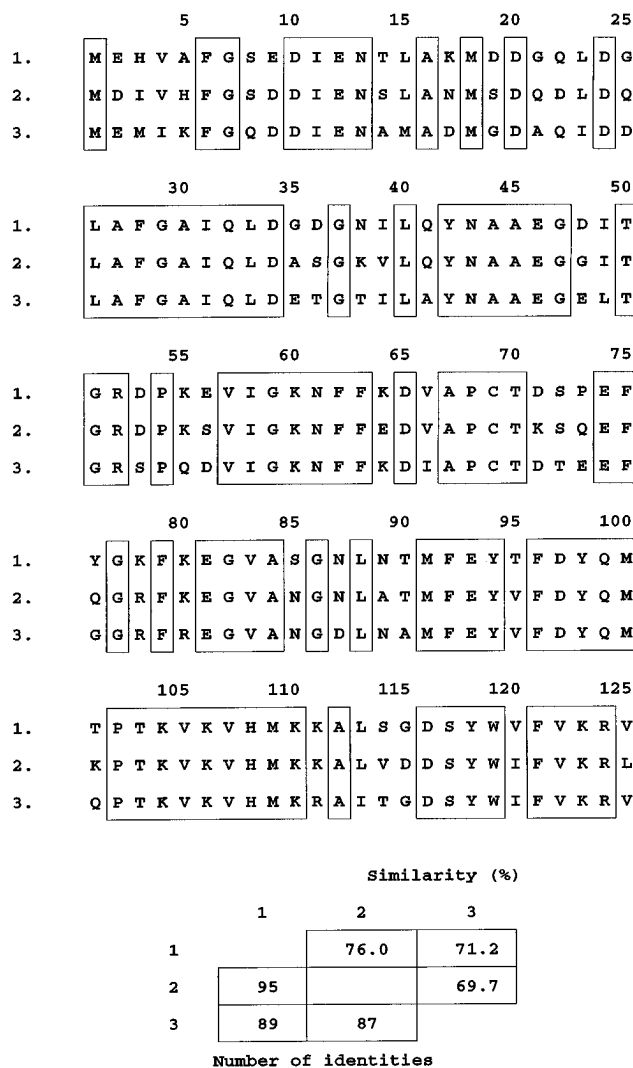


FIGURE 8: Alignment of amino acid sequences of the photoactive yellow proteins from *E. halophila* (1), *C. salexigens* (2), and *R. salexigens* (3). Identical residues are boxed. Cys69 is the site of covalent chromophore attachment. The number of identical residues from this alignment and the calculated similarities are presented in the lower part of the figure. Note that all but one of the residues identified by Borgstahl et al. (1995) as interacting with the chromophore are conserved (Ile31, Tyr42, Glu46, Thr50, Arg52, Phe62, Ala67, Phe96, and Tyr98) except Val66, which is Ile in RSPYP.

different requirement for structural stability rather than to a crossing-over event which may also arise if more than one gene for PYP were present. However, we have no indication from the protein sequence data that there is more than one copy of PYP per genome. The EHPYP gene has been cloned and there is no evidence for more than one copy per genome either (Baca et al., 1994). The chromophore is bound to the C-terminal half of the protein and is separated from the N-terminal segment by the central  $\beta$ -sheet (Borgstahl et al., 1995). The relatively greater distance to the active site may in fact be the reason for the greater variability of the N-terminal region.

It is perhaps not surprising that PYP is highly conserved if one considers that a transient protein conformational change occurs during the *E. halophila* photocycle and that a hydrophobic site is exposed to solvent (Meyer et al., 1989). The refined three-dimensional structure suggests that the chromophore might be exposed to solvent when it isomerizes about the acrylic double bond which is responsible for the

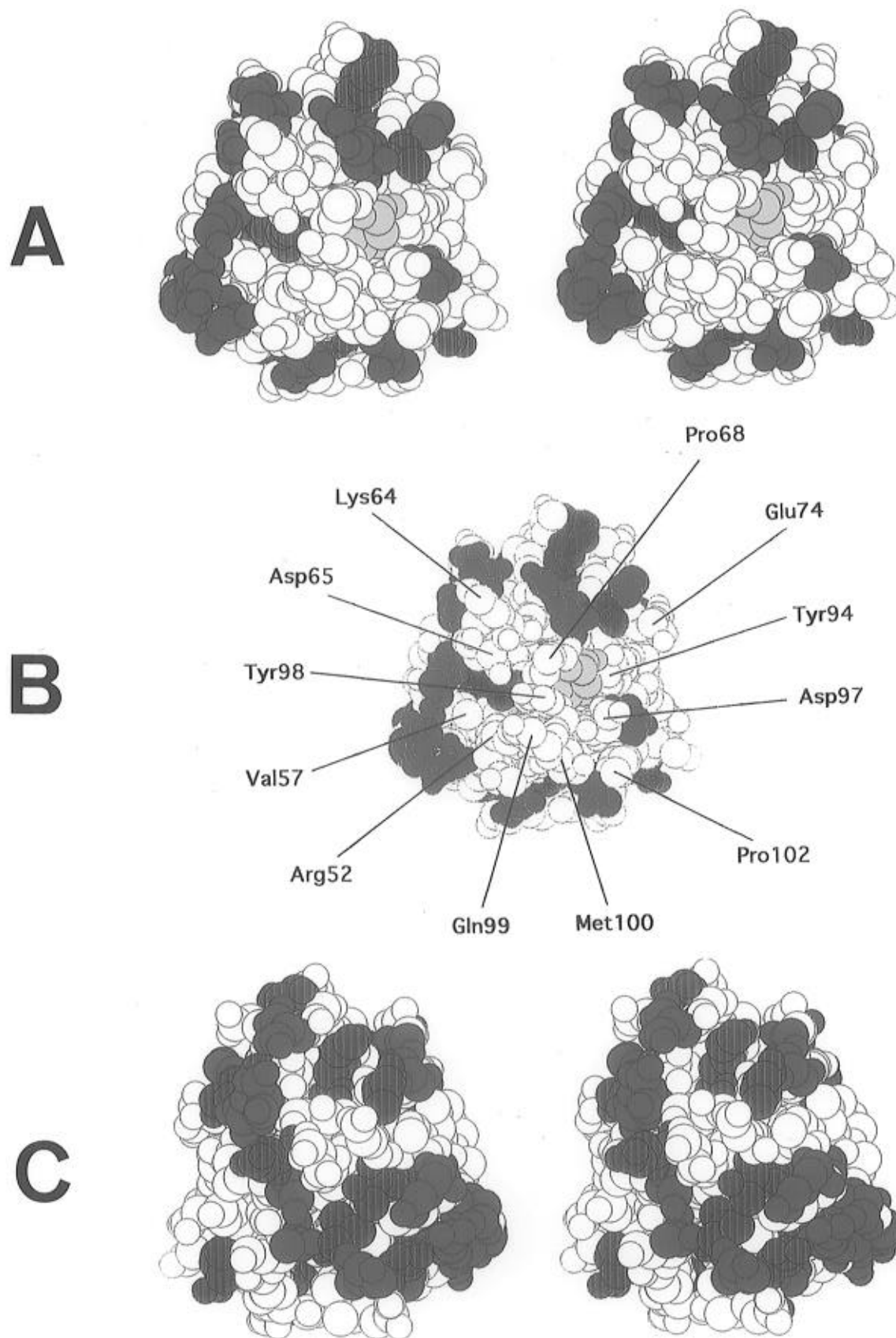


FIGURE 9: (A) Stereoview of the three-dimensional structure of *E. halophila* PYP [after Borgstahl et al. (1995)] showing the spatial location of invariant residues (white), conservative substitutions (green), and radical changes (red) occurring in CSPYP and RSPYP relative to EHPYP. The yellow region represents the chromophore bound to Cys69. The invariant surface near the chromophore involves the residues shown in panel B: Asp65, Lys64, Pro68, Cys69, Glu74, Tyr94, Asp97, Pro102, Met100, Gln99, Tyr98, Arg52, and Val57. (C) Stereoview after 180° rotation around the Y-axis.

photoactivity, and that Tyr98 and perhaps Val66 and Arg52 would have to swing out of its way (Borgstahl et al., 1995).

Thus, the hydrophobic residues which become more exposed during the photocycle are probably the chromophore itself,



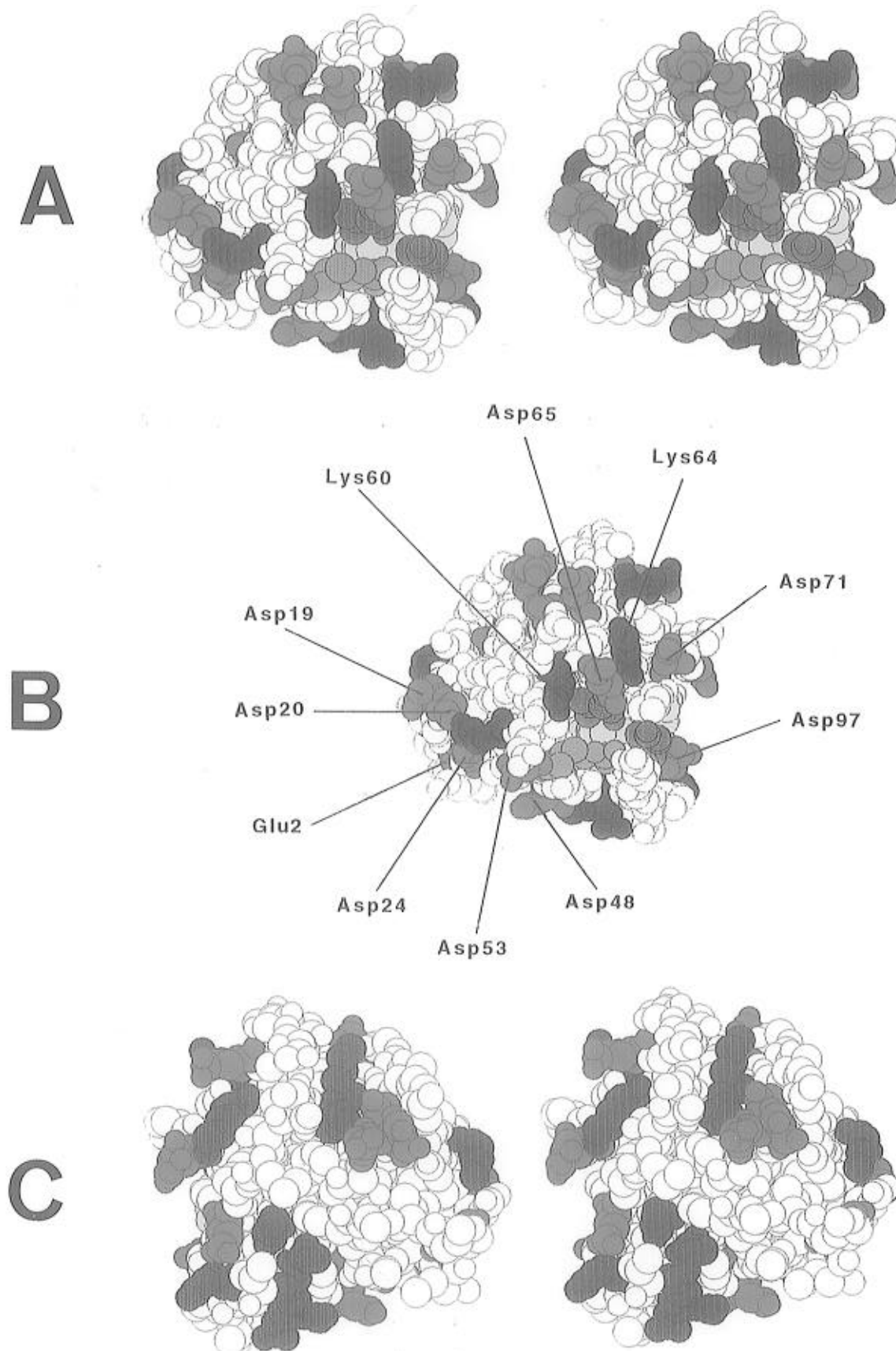


FIGURE 10: (A) Stereoview of *E. halophila* PYP showing basic residues (blue and cyan), acidic residues (red), Cys69 and the chromophore (yellow), Arg52 (cyan), Val66 (green, above chromophore), and Tyr 98 (green, to the right of the chromophore). The negative patch discussed by Borgstahl et al. (1995) is composed of the residues shown in panel B. As shown in Figure 8, it is only partially conserved. Nearest charges above the chromophore are Lys60, Asp65, Lys64, and Asp71. Only Lys64 and Asp65 are conserved. (C) Stereoview after 180° rotation around the Y-axis.

Val66 and Tyr98. It was also found that *E. halophila* PYP transiently takes up a proton during the photocycle (Meyer

et al., 1993). The chromophore hydroxyl is ionized in the ground state and stabilized by the nearby charge on Arg52

(Borgstahl et al., 1995). The initial step in the photocycle is cis–trans isomerization about the acrylic double bond which then exposes the chromophore to the solvent. Thus, the chromophore itself may be responsible for the proton uptake during the dark reaction, or second phase, of the photocycle. Proton release occurs during the relaxation, or third phase, back to the resting state.

Although it may only be a coincidence, we have discovered that one of the two proteins which accompany *C. salexigens* PYP up to the last purification step (Figure S.7) is a homolog of the *E. coli* and *Salmonella typhimurium* “CheY” chemotaxis protein (Stock et al., 1985, 1989; Mutoh & Simon, 1986). As a working hypothesis, we suggest that photoactivated PYP transiently binds to the flagellar motor by analogy with phosphorylated CheY. An intracellular location for PYP would be necessary for interaction with the flagellar motor in this manner. The gene sequence shows that the N-terminal methionine in PYP is the initiator methionine and that PYP is in fact cytoplasmic (Baca et al., 1994).

In order to locate a possible site of interaction of PYP with other proteins such as the flagellar motor, we have mapped, as described by Meyer et al. (1994), the amino acid substitutions of the new sequences onto the three-dimensional structure of *E. halophila* PYP (Borgstahl et al., 1995). The result is given in Figure 9. There is in fact a conserved surface in PYP near the Cys69 chromophore binding site and centered over Tyr98. This is as expected if the conformational change exposing the chromophore is important to the function and necessary for binding to other proteins. We propose that this conserved surface is the site of interaction with reaction partner proteins. This site may be further delineated via additional protein sequences and site-directed mutations.

Surface charges are often involved in stabilizing protein–protein complexes, especially when complementary positive and negative charges are localized at the two sites of interaction of the reaction partners (Figure 10). Borgstahl et al. (1995) identified a region of localized negative charge in EHPYP. However, this negatively charged site is not very close to Tyr98 and is one of the least conserved regions of the protein surface. Nevertheless, the negative surface is even more highly charged in the new sequences. Although there could be some unknown significance to this concentration of charge, we believe that is has no direct functional role. The conserved surface near Tyr98 has a more balanced charge distribution and such charge may have a minimal role in binding reaction partners. We believe that surface shape complementarity in the bleached state is most important for binding the postulated protein receptor via hydrophobic forces.

## SUPPORTING INFORMATION AVAILABLE

Six figures showing RP-HPLC chromatograms of digested protein and a plot of released C-terminal amino acids vs time after treatment with carboxypeptidases A and B and 13 tables listing sequence and mass analysis data for various peptides (28 pages). Ordering information is given on any current masthead page.

## REFERENCES

Ambler, R. P., Daniel, M., Hermoso, J., Meyer, T. E., Bartsch, R. G., & Kamen, M. D. (1979) *Nature* 278, 659–660.

- Ambler, R. P., Meyer, T. E., & Kamen, M. D. (1994a) *Arch. Biochem. Biophys.* 308, 78–81.
- Ambler, R. P., Daniel, M., Meyer, T. E., & Kamen, M. D. (1994b) *Biochimie* 76, 583–591.
- Baca, M., Borgstahl, G. E. O., Biossinot, M., Burke, P. M., Williams, D. R., Slater, K. A., & Getzoff, E. D. (1994) *Biochemistry* 33, 14369–14377.
- Biebl, H., & Pfenning, N. (1981) in *The Prokaryotes. A Handbook on Habitats, Isolation and Identification of Bacteria* (Starr, M. P., Stolp, H., Ballows, A., & Schlegel, H. G., Eds.) pp 267–273, Springer-Verlag, Berlin, Germany.
- Bogomolni, R. A., & Spudich, J. L. (1991) in *Sensory Receptors and Signal Transduction*, pp 233–255, Wiley-Liss, Inc., New York.
- Borgstahl, G. E. O., Williams, D. R., & Getzoff, E. D. (1995) *Biochemistry* 34, 6278–6287.
- Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem.* 231, 622.
- Drews, G. (1981) *Arch. Microbiol.* 130, 325–327.
- Hoff, W. D., Kwa, S. L. S., Van Grondelle, R., & Hellingwerf, K. J. (1992) *Photochem. Photobiol.* 56, 529–539.
- Hoff, W. D., Düx, P., Hård, Devreese, B., K., Nugteren-Roodzant, I. M., Crielaard, W., Boelens, R., Kaptein, R., Van Beeumen, J., & Hellingwerf, K. J. (1994a) *Biochemistry* 33, 13959–13962.
- Hoff, W. D., Van Stokkum, I. H. M., Van Ramesdonk, H. J., Van Brederode, R., Brouwer, A. M., Fitch, J. C., Meyer, T. E., Van Grondelle, R., & Hellingwerf, K. J. (1994b) *Biophys. J.* 67, 1691–1705.
- Laemmli, U. K. (1970) *Nature* 227, 680–685.
- Mandel, M., Leadbetter, E. R., Pfenning, N., & Trüper, H. G. (1971) *Int. J. Syst. Bacteriol.* 21, 222–230.
- McRee, D. E., Tainer, J. A., Meyer, T. E., Van Beeumen, J., Cusanovich, M. A., & Getzoff, E. D. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 6533–6537.
- Meyer, T. E. (1985) *Biochim. Biophys. Acta* 806, 175–183.
- Meyer, T. E., Yakali, E., Cusanovich, M. A., & Tollin, G. (1987) *Biochemistry* 26, 418–423.
- Meyer, T. E., Tollin, G., Hazzard, J. H., & Cusanovich, M. A. (1989) *Biophys. J.* 56, 559–564.
- Meyer, T. E., Fitch, J. C., Bartsch, R. G., Tollin, G., & Cusanovich, M. A. (1990) *Biochim. Biophys. Acta* 1016, 364–370.
- Meyer, T. E., Tollin, G., Causgrove, T. P., Cheng, P., & Blankenship, R. E. (1991) *Biophys. J.* 59, 988–991.
- Meyer, T. E., Cusanovich, M. A., & Tollin, G. (1993) *Arch. Biochem. Biophys.* 306, 515–517.
- Meyer, T. E., Tollin, G., & Cusanovich, M. A. (1994) *Biochimie* 76, 480–488.
- Miller, A., Leigeb, H., Hoff, W. D., & Hellingwerf, K. J. (1993) *Biochim. Biophys. Acta* 1141, 190–196.
- Moore, G. R., & Pettigrew, G. W. (1990) in *Cytochromes C. Evolutionary, Structural and Physicochemical Aspects*, pp 133–137, Springer Verlag, Berlin, Germany.
- Mutoh, N., & Simon, M. I. (1986) *J. Bacteriol.* 165, 161–166.
- Ng, K., Getzoff, E. D., & Moffat, K. (1995) *Biochemistry* 34, 879–890.
- Pfenning, N., & Trüper, H. G. (1981) in *The Prokaryotes. A Handbook on Habitats, Isolation and Identification of Bacteria* (Starr, M. P., Stolp, H., Ballows, A., & Schlegel, H. G., Eds.) pp 279–289, Springer-Verlag, Berlin, Germany.
- Stock, A. M., Koshland, D. E., Jr., & Stock, J. B. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 7989–7993.
- Stock, A. M., Mottonen, J. M., Stock, J. B., & Schutt, C. E. (1989) *Nature* 337, 745–749.
- Trüper, H. G., & Imhoff, R. (1981) in *The Prokaryotes. A Handbook on Habitats, Isolation and Identification of Bacteria* (Starr, M. P., Stolp, H., Ballows, A., & Schlegel, H. G., Eds.) pp 274–278, Springer-Verlag, Berlin, Germany.
- Van Beeumen, J. J., Devreese, B. V., Van Bun, S. M., Hoff, W. D., Hellingwerf, K. J., Meyer, T. E., McRee, D. E., & Cusanovich, M. A. (1993) *Protein Sci.* 2, 1114–1125.
- Van Brederode, M. E., Gensch, T., Hoff, W. D., Hellingwerf, K. J., & Braslavsky, S. E. (1995) *Biophys. J.* 68, 1101–1109.

BI951494T