Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Short hydrogen bonds in photoactive yellow protein

Spencer Anderson,[a,b]* Sean Crosson[b] and Keith Moffat[a,b,c]

[a]Consortium for Advanced Radiation Sources, University of Chicago, USA, [b]Department of Biochemistry and Molecular Biology, University of Chicago, USA, and [c]Institute for Biophysical Dynamics, University of Chicago, USA

Correspondence e-mail: smander@midway.uchicago.edu

Received 9 January 2004
Accepted 15 March 2004

PDB References: WT PYP, $P6_3$, 295 K, 1otb, r1otbsf; E46Q PYP, $P6_3$, 295 K, 1ota, r1otasf; WT PYP, $P6_3$, 110 K, 1ot9, r1ot9sf; E46Q PYP, $P6_3$, 110 K, 1ot6, r1ot6sf; E46Q PYP, $P6_5$, 295 K, 1oti, r1otisf; E46Q PYP, $P6_5$, 110 K, 1ote, r1otesf; WT PYP, $P2_1$, 110 K, 1otd, r1otdsf.

Eight high-resolution crystal structures of the ground state of photoactive yellow protein (PYP) solved under a variety of conditions reveal that its chromophore is stabilized by two unusually short hydrogen bonds. Both Tyr42 O^{η} and Glu46 O^{ε} are separated from the chromophore phenolate oxygen by less than the sum of their atomic van der Waals radii, 2.6 Å. This is characteristic of strong hydrogen bonding, in which hydrogen bonds acquire significant covalent character. The hydrogen bond from the protonated Glu46 to the negatively charged phenolate oxygen is 2.58 ± 0.01 Å in length, while that from Tyr42 is considerably shorter, 2.49 ± 0.01 Å. The E46Q mutant was solved to 0.95 Å resolution; the isosteric mutation increased the length of the hydrogen bond from Glx46 to the chromophore by 0.29 ± 0.01 Å to that of an average hydrogen bond, 2.88 ± 0.01 Å. The very short hydrogen bond from Tyr42 explains why mutating this residue has such a severe effect on the ground-state structure and PYP photocycle. The effect of isosteric mutations on the photocycle can be largely explained by the alterations to the length and strength of these hydrogen bonds.

1. Introduction

PYP is a soluble 14 kDa blue-light photoreceptor that was initially isolated from *Halorhodospira halophila* (Meyer, 1985). PYP belongs to the PAS-domain superfamily, which has been widely implicated in signaling processes in many species (Taylor & Zhulin, 1999). The visible absorption spectrum of PYP is very similar to the wavelength-dependence of phototaxis in *H. halophila*, which suggests that PYP is involved in the negative phototactic response to blue light (Sprenger *et al.*, 1993). PYP absorbs blue light *via* its *para*-coumaric acid chromophore (pCA), which is covalently attached to Cys69. Photoactivated PYP enters a fully reversible photocycle that ultimately transduces the photon energy into a biologically important structural change (Cusanovich & Meyer, 2003; Hellingwerf *et al.*, 2003).

The ground-state structure of wild-type PYP has been solved previously in two crystallographic space groups and in solution using NMR (Borgstahl *et al.*, 1995; Dux *et al.*, 1998; Genick *et al.*, 1998; van Aalten *et al.*, 2000). The ground-state chromophore is in a *trans* conformation in which the negatively charged chromophore phenolate oxygen participates in a hydrogen-bond network with Tyr42, Glu46 and Thr50 (Fig. 1) (Borgstahl *et al.*, 1995). The properties of this hydrogen-bond network are critical to establishing the ground-state stability, absorbance maximum and photocycle kinetics (Devanathan *et al.*, 1999; Genick, Devanathan *et al.*, 1997; Kroon *et al.*, 1996). The network is modified in the E46Q mutant of PYP. This isosteric mutant is proposed to have a weakened hydrogen

bond from Gln46 $N^{\varepsilon 2}$ to the chromophore phenolate oxygen, which causes greater negative electron density on the chromophore and accounts for the red shift in its absorbance maximum from 446 nm in wild type to 462 nm in E46Q (Genick, Devanathan *et al.*, 1997).

The wild-type chromophore-binding pocket is unusual in that the pK_a values of several groups are substantially shifted by both hydrogen bonding and desolvation effects (Harris & Turner, 2002). At neutral pH, the chromophore phenolate oxygen is deprotonated (Kim *et al.*, 1995) and the Glu46 side chain is protonated (Xie *et al.*, 1996). The Gln46 side chain in the E46Q mutant can neither donate a proton to the chromophore nor acquire a negative charge during the photocycle, key roles proposed for Glu46 in wild-type PYP (Imamoto *et al.*, 1997; Xie *et al.*, 1996). Although the mutant has a faster photocycle, every spectroscopic intermediate found in the photocycle of wild-type PYP has a direct counterpart in the E46Q mutant (Devanathan *et al.*, 2000; Genick, Devanathan *et al.*, 1997; Zhou *et al.*, 2001).

We have analyzed crystallographically the effects of the conservative isosteric E46Q mutation, temperature, crystallization solvent and space group on the hydrogen-bond network to the chromophore. Since we are interested in accurate bond lengths and small structural differences, it was essential to obtain crystallographic data with extremely high resolution and use comparable refinement strategies and constraints on all structures. We identify unique aspects of the hydrogen-bond network in both wild-type and mutant PYP and relate changes in hydrogen-bond strengths to the lifetimes of photocycle intermediates.

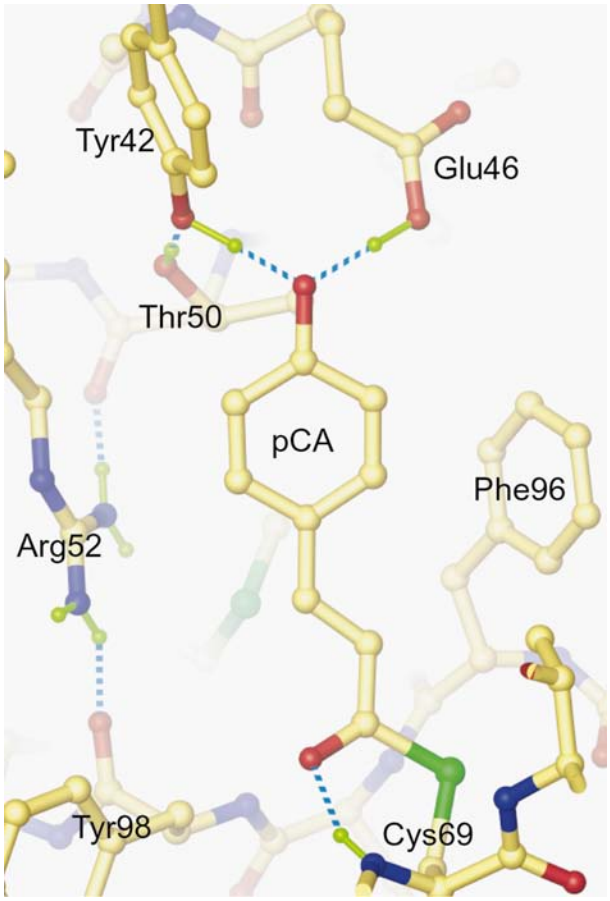

Figure 1
The ground-state coumaric acid chromophore (pCA) and its binding pocket in wild-type PYP. Hydrogen bonds are shown as dashed lines.

2. Materials and methods

2.1. Protein preparation

The wild-type *H. halophila* (strain SL-1) sample was initially cloned from a bacterial culture (DSM, Braunschweig, Germany). All protein samples were prepared by heterologous overexpression of histidine-tagged apoprotein (Kort *et al.*, 1996). After affinity purification of wild-type and E46Q mutant protein, the polyhistidine affinity tag was removed by incubation with enterokinase (strain BN9626) or TEV protease (strain SL-1) at room temperature. Sample purity was assessed both by the presence of a single band on an SDS polyacrylamide gel and by optical purity (the ratio of optical absorbance at 280 nm to that at the chromophore peak: 446 nm for wild-type PYP and 462 nm for the E46Q mutant). The optical purity was 0.44 for wild-type PYP (strain BN9626), 0.41 for wild-type PYP (strain SL-1) and 0.42 for E46Q mutant PYP (strain BN9626).

2.2. Crystallization

Wild-type and E46Q mutant *H. halophila* (strain BN9626) protein were concentrated to 30 mg ml^{-1} and then crystallized in space group $P6_3$ using microseeding at 293 K in 2.6 *M* ammonium sulfate, 50 m*M* sodium phosphate pH 7.0 (McRee *et al.*, 1989). The E46Q mutant was also crystallized in the space group $P6_5$ at 293 K in 40% PEG 2K, 50 m*M* MES pH 6.5 (van Aalten *et al.*, 2000). PYP from *H. halophila* (strain SL-1) was concentrated to 30 mg ml^{-1} in 100 m*M* Tris pH 8.0 and crystallized in 35% PEG 4000, 350 m*M* $MgCl_2$, 100 m*M* Tris. Large crystals in space group $P2_1$ only grew at pH 9.5 or higher.

2.3. Data collection

For data collection at 110 K, crystals in space groups $P6_5$ and $P2_1$ were frozen directly in mother liquor, while 15% glycerol was added as a cryoprotectant when freezing crystals in space group $P6_3$. For data collection at 295 K, crystals were transferred to stabilization buffer and then mounted in 700 µm diameter glass capillaries. The stabilization buffer was mother liquor for crystals in space group $P6_5$ and 3.2 *M* ammonium sulfate, 50 m*M* sodium phosphate buffer pH 7.0 for crystals in space group $P6_3$.

All diffraction data were collected using the monochromatic oscillation technique at the Advanced Photon Source BioCARS beamline 14BM-C, Argonne National Laboratory using an ADSC Quantum 4 detector. 1.0 Å wavelength X-rays were used for most data collections, except for the two E46Q mutant crystals in space group $P6_5$ and the wild-type (strain BN9626) data at 295 K, which utilized 0.9 Å X-rays.

2.4. Data processing and refinement

DENZO and *SCALEPACK* (Otwinowski & Minor, 1997) were used for indexing, integration, scaling and merging of monochromatic oscillation data. Initial phases were determined using molecular replacement with the wild-type PYP structure (PDB code 2phy; Borgstahl *et al.*, 1995) as a search model in *CNS* (Brünger *et al.*, 1998). Model building was performed using *XtalView* (McRee, 1999) and *SHELX*97 (Sheldrick, 1997) was used for all refinement. Later stages of all refinement used anisotropic *B* values, excluding the relatively disordered three N-terminal residues of each model. Although the data in space group $P6_5$ only extend to 1.4 Å resolution, $R_{\rm free}$ immediately fell by >4% when anisotropic *B* values were introduced. H atoms were not refined for any of the structures; no H atoms involved in the hydrogen-bond network to the chromophore were apparent in difference electron-density maps.

Chromophore bond-distance and angle restraints were necessary for the initial stages of refinement of the four $P6_3$ structures; these restraints were generated from the 2phy wild-type PYP structure. Restraints for the chromophore, Tyr42, Glx46 and Thr50 were removed at the later stages of the refinement since the electron density in this region was very well ordered. Chromophore restraints generated from our refined structures in space group $P6_3$ were retained throughout the lower resolution refinements of the 1.25 Å $P2_1$ wild-type (strain SL-1) and 1.4 Å $P6_5$ E46Q mutant structures. All structures were refined to convergence using conjugate-gradient least-squares refinement. We performed a single round of block-matrix least-squares refinement on the four atomic resolution structures in space group $P6_3$ in order to calculate estimated standard uncertainties (Sheldrick, 1997). The 125-amino-acid molecule was divided into three overlapping blocks of 50 residues each; all residues involved in the hydrogen-bond network to the chromophore were included in a single block. These estimated standard deviations calculated from block-matrix refinement are slightly underestimated since the refinement does not consider covariance between parameters in the different blocks.

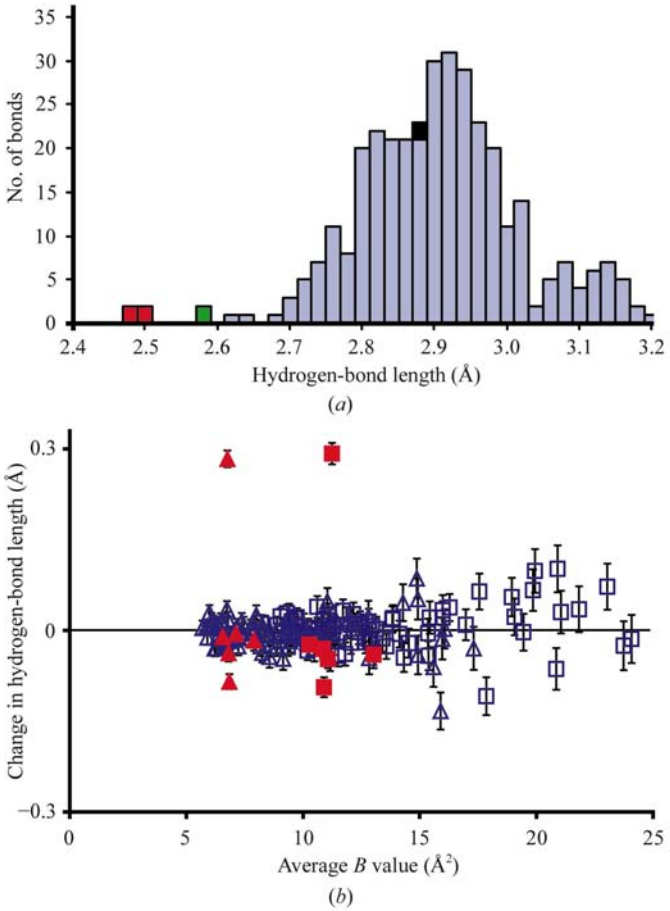

Figure 2
Hydrogen bonds in the wild-type and E46Q structures in space group $P6_3$ at 110 and 295 K. (*a*) Distribution of hydrogen-bond lengths in the four structures. Bonds to the phenolate oxygen of the chromophore from Tyr42 are identified in red, from Glu46 in wild type in green and from Gln46 in the E46Q mutant in black. (*b*) Difference in hydrogen-bond lengths between the E46Q mutant and wild type. Open triangles denote difference in bond lengths at 110 K and open squares that at 295 K. Solid red symbols denote the six hydrogen bonds in the chromophore-binding pocket (Fig. 1 and Table 2).

3. Results

The ground-state crystal structures of both wild-type PYP and the E46Q mutant were solved under multiple conditions. A total of seven structures (containing eight PYP monomers) were determined to resolutions between 0.95 and 1.40 Å (Table 1). The highest resolution data to between 0.95 and 1.10 Å was collected on crystals belonging to space group $P6_3$; data collected on other crystal forms was used to validate results obtained from the $P6_3$ crystals.

3.1. Wild-type PYP structures

We solved and refined the structure of wild-type PYP (strain BN9626) in space group $P6_3$ at 295 and 110 K. Both structures display unusually short hydrogen bonds from Tyr42 and Glu46 to the chromophore phenolate oxygen (Table 2). The hydrogen bond from Tyr42 to the phenolate oxygen is 2.51 ± 0.01 Å and 2.50 ± 0.01 Å at 295 and 110 K, respectively. The unusually short length of this hydrogen bond clearly classifies it as a 'strong' or 'low-barrier' hydrogen bond (Cleland, 2000; Hibbert & Emsley, 1990; Steiner, 2002). The length of the hydrogen bond from the protonated Glu46 to the chromophore phenolate oxygen is 2.58 ± 0.01 and 2.59 ± 0.01 Å at 295 and 110 K, respectively, which is slightly less than the sum of the van der Waals radii of the O atoms. Although longer than the hydrogen bond from Tyr42, this hydrogen bond is considerably shorter than an average hydrogen bond in PYP and is still classified as a strong hydrogen bond (Fig. 2*a*).

The novel wild-type PYP structure from *H. halophila* (strain SL-1) was also determined in space group $P2_1$ to 1.25 Å resolution at 110 K. Contrary to a report that the sequence of PYP (strain SL-1) is identical to that of PYP (strain BN9626; Kort *et al.*, 1996), they differ at six residues (I. Dzanibekova, B. Perman & K. Moffat, unpublished work). These differences

Table 1
Crystallographic data and refinement statistics for wild-type (WT) PYP and the E46Q mutant.

Values in parentheses refer to the last shell.

	WT	E46Q	WT	E46Q	E46Q	E46Q	WT (SL-1)
Space group	$P6_3$	$P6_3$	$P6_3$	$P6_3$	$P6_5$	$P6_5$	$P2_1$
Temperature (K)	295	295	110	110	295	110	110
Resolution (Å)	1.10	1.10	1.00	0.95	1.40	1.40	1.25
Unit-cell parameters (Å)							
$a = b$	66.83	66.92	66.21	66.19	40.99	40.02	†
c	40.95	41.00	40.51	40.50	117.28	117.61	†
Total observations	721991	414335	860006	855706	495366	513415	389278
Unique reflections	42402	42583	54493	63595	21402	20725	61494
Redundancy	17.0	9.7	15.8	13.5	23.1	24.8	6.3
Resolution (Å)	100–1.10 (1.14–1.10)	100–1.10 (1.14–1.10)	100–1.00 (1.04–1.00)	100–0.95 (0.98–0.95)	100–1.40 (1.45–1.40)	100–1.40 (1.45–1.40)	100–1.25 (1.29–1.25)
Completeness (%)	96.8 (81.5)	89.0 (71.7)	98.6 (95.4)	93.5 (78.5)	95.4 (96.8)	96.5 (94.3)	90.3 (56.9)
$I/\sigma(I)$	28.6 (2.5)	20.0 (4.8)	30.8 (7.2)‡	26.5 (4.4)	12.4 (2.6)	22.3 (13.2)	16.3 (2.0)
$R_{\rm merge}$§	5.3 (53.2)	5.1 (20.0)	5.7 (56.8)‡	4.6 (25.4)	8.8 (56.8)	7.0 (13.0)	6.1 (40.0)
$R_{\rm cryst}$¶	13.26	15.28	13.39	14.37	13.95	18.47	17.06
$R_{\rm free}$††	16.06	18.02	16.14	16.11	18.42	22.82	21.47
No. ordered waters	139	138	195	185	117	138	224
No. atoms	1126	1125	1182	1172	1104	1125	1908
No. parameters	10137	10128	10641	10551	9939	10128	19785
No. restraints	11374	11373	11405	11403	11368	11373	21091
Average B value	19.4	18.9	12.3	13.1	27.0	33.5	24.6
R.m.s.d. bond lengths (Å)	0.015	0.016	0.016	0.016	0.014	0.015	0.016
R.m.s.d. bond angles (°)	2.31	2.44	2.29	2.40	2.55	2.44	2.61
Ramachandran distribution							
Most favored (%)	90.6	89.6	88.7	91.5	87.7	86.4	90.4
Allowed (%)	9.4	10.4	11.3	8.5	12.3	13.6	9.6
PDB code	1otb	1ota	1ot9	1ot6	1oti	1ote	1otd

† $P2_1$ unit-cell parameters are $a = 33.98$, $b = 91.10$, $c = 36.76$ Å, $\alpha = 90$, $\beta = 92.7$, $\gamma = 90°$. ‡ All data were collected on a single crystal but on two separate occasions because of limitations in beamtime. The high $R_{\rm merge}$ is likely to be a result of damage from radiation or refreezing the crystal. § $R_{\rm merge} = \sum_{hkl}\sum_i |I_i - \langle I\rangle|/\sum_{hkl}\sum_i I_i$ for all data. ¶ $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$, including all data. †† $R_{\rm free}$ uses 10% of data for the test set.

Table 2
Length of hydrogen bonds (Å).

Lengths less than 2.65 Å are in bold. Estimated standard deviations of bond lengths are shown for the structures in space group $P6_3$. NA, not applicable. PDB codes: 3pyp (Genick *et al.*, 1998), 2phy (Borgstahl *et al.*, 1995), 1f98 (Brudler *et al.*, 2000), 1f9i (Brudler *et al.*, 2000), 1d7e (van Aalten *et al.*, 2000).

Structure	Temperature (K)	Resolution (Å)	Glx46 O/N$^{\varepsilon 2}$ to pCA phenolate O	Tyr42 O$^{\eta}$ to pCA phenolate O	Thr50 O$^{\gamma}$ to Tyr42 O$^{\eta}$	Arg52 N$^{\eta 1}$ to Thr50 O	Arg52 N$^{\eta 2}$ to Tyr98 O	Cys69 N to pCA O^{1}
$P6_3$ structures								
E46Q	110	0.95	2.88 ± 0.01	**2.48 ± 0.01**	2.78 ± 0.01	2.88 ± 0.01	2.92 ± 0.01	2.74 ± 0.01
WT	110	1.00	**2.59 ± 0.01**	**2.50 ± 0.01**	2.85 ± 0.01	2.93 ± 0.01	2.93 ± 0.01	2.75 ± 0.01
Difference			0.29 ± 0.01	−0.02 ± 0.01	−0.07 ± 0.01	−0.05 ± 0.02	−0.01 ± 0.02	−0.01 ± 0.01
E46Q	295	1.10	2.87 ± 0.01	**2.49 ± 0.01**	2.84 ± 0.01	2.89 ± 0.02	2.90 ± 0.02	2.80 ± 0.01
WT	295	1.10	**2.58 ± 0.01**	**2.51 ± 0.01**	2.89 ± 0.01	2.93 ± 0.02	2.93 ± 0.02	2.82 ± 0.01
Difference			0.30 ± 0.02	−0.02 ± 0.02	−0.05 ± 0.01	−0.04 ± 0.02	−0.03 ± 0.02	−0.02 ± 0.02
WT (3pyp)	149	0.82	**2.61**	**2.49**	2.81	2.93	3.00	2.76
WT (2phy)	295	1.40	2.69	2.71	2.83	2.97	2.92	2.69
T50V (1f98)	100	1.15	**2.58**	**2.62**	NA	2.80	2.91	2.81
Y42F (1f9i)	100	1.10	**2.51**	NA	2.79 to pCA	2.90	2.83	2.70
$P6_5$ structures								
E46Q	110	1.40	2.89	**2.47**	2.83	2.89	2.75	2.84
E46Q	295	1.40	2.93	**2.49**	2.88	2.88	2.85	2.78
WT (1d7e)	100	1.39	**2.51**	**2.50**	2.80	2.93	2.81	2.81
$P2_1$ structure								
WT '*A*' chain	110	1.25	**2.54**	**2.47**	2.91	2.92	2.86	2.69
WT '*B*' chain	110	1.25	**2.56**	**2.48**	2.81	2.99	2.84	2.77

do not alter the absorbance maximum (data not shown) or protein fold (Fig. 3*a*) since all are located on the surface of the molecule: G21S, G25N, N38T, D53N, Q56E and Y76S. Although the N-terminal cap is largely disordered in both monomers of the crystallographic dimer (Fig. 4), the structure of the protein core is very well defined. Both molecules in the asymmetric unit exhibit strong hydrogen bonds from Tyr42 to the chromophore phenolate oxygen, with lengths of 2.47 and 2.49 Å. The length of the hydrogen bond from the protonated Glu46 to the phenolate oxygen is 2.53 and 2.55 Å, slightly shorter than that in space group $P6_3$.

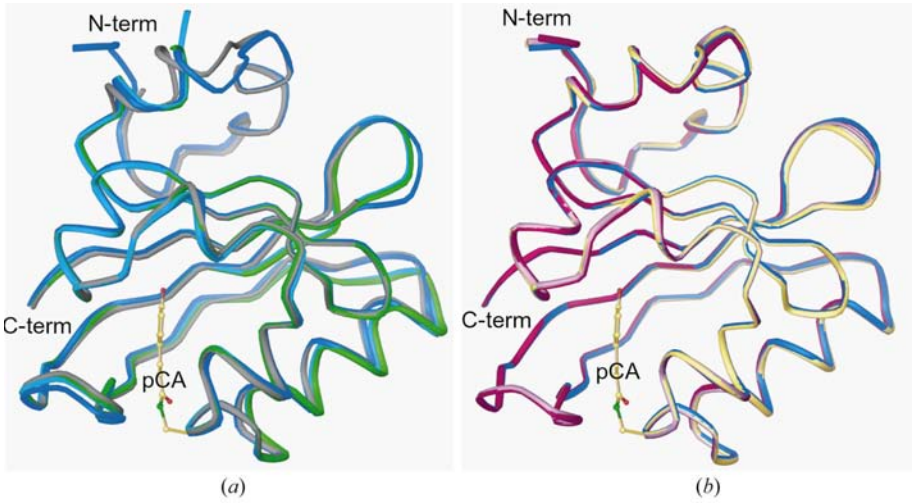

Figure 3
Alignment of the PYP structures. All structures were aligned on the coumaric acid chromophore and Cys69. (*a*) Wild-type PYP structures at cryogenic temperature in three different space groups: space group $P6_3$ (dark blue), space group $P2_1$ (chain *A*, light blue; chain *B*, green) and space group $P6_5$ (PDB code 1d7e; van Aalten *et al.*, 2000) (grey). (*b*) Wild-type and E46Q mutant structures in space group $P6_3$ at two temperatures: wild type at 295 K, yellow; E46Q at 295 K, pink; wild type at 110 K,dark blue; E46Q at 110 K, purple. The r.m.s.d.s (Å) from the alignment of all atoms (C^{α} atoms only) between wild type at 110 K and E46Q at 110 K is 0.54 (0.14), between wild type at 295 K and E46Q at 295 K is 0.62 (0.14), between wild type at 110 K and wild type at 295 K is 0.71 (0.29) and between E46Q at 110 K and E46Q at 295 K is 0.70 (0.27).

3.2. E46Q mutant structures

The E46Q mutant was crystallized in space group $P6_3$ and refined to 1.10 Å resolution at 295 K and to 0.95 Å at 110 K (Table 1). As expected, the isosteric replacement of a protonated glutamic acid with glutamine does not alter the global protein structure (Fig. 3*b*); all significant structural differences are restricted to the chromophore-binding pocket (Figs. 5*a* and 5*b*). The orientation of the Gln46 side chain could be established since oxygen is clearly distinguishable from nitrogen at this high resolution. Gln46 $N^{\varepsilon 2}$ acts as the hydrogen-bond donor to the chromophore in the E46Q mutant. The E46Q mutation has a smaller effect than either temperature or space group on the overall structure as evidenced by the r.m.s.d. alignment of all atoms (Fig. 3*b*).

As in wild-type PYP, hydrogen-bond lengths in the E46Q mutant are unaffected by temperature. Only those bonds directly involved in the hydrogen-bond network to the chromophore have significantly different lengths (Fig. 2*b*). The largest structural change caused by the E46Q mutation is an increase of 0.29 ± 0.01 Å at 110 K and 0.30 ± 0.02 Å at 295 K in the length of the short hydrogen bond from Glx46 to the chromophore. The length of this hydrogen bond in the mutant reverts to that of an average hydrogen bond (Table 2; Fig. 2*a*). In contrast, the unusually short hydrogen bond from Tyr42 to the chromophore phenolate oxygen is retained in the mutant, with lengths of 2.48 ± 0.01 Å at 110 K and 2.49 ± 0.01 Å at 295 K. The chromophore is displaced slightly out of its binding pocket towards solvent in order to accommodate the longer hydrogen bond from Gln46 (Figs. 5*c* and 5*d*). Tyr42 and Thr50 are both displaced along with the chromophore, thus preserving the length of the short hydrogen bond from Tyr42 to the chromophore phenolate oxygen. The hydrogen bond from Thr50 to Tyr42 is slightly shorter in the E46Q mutant at both 295 and 110 K; although this length change is minor, its magnitude is greater than twice the estimated standard deviation of this value (Table 2).

We also solved the E46Q mutant structure in space group $P6_5$ to 1.4 Å resolution at both 110 and 295 K (Table 1). The hydrogen-bond lengths in the chromophore-binding pocket are very similar to those in the space group $P6_3$ structure (Table 2). There is continuous electron density for the loop from residues 114 to 116 at both temperatures, in contrast to the wild-type structure in this space group in which these atoms are disordered (van Aalten *et al.*, 2000). The mosaicity in the $P6_5$ crystal at 110 K is high, which lowered the quality of the data, as evidenced by the higher *R* and *B* values (Table 1). However, the refined hydrogen-bond lengths in the chromophore-binding pocket are nearly identical to those measured in the higher quality $P6_5$ data set at 295 K (Table 2).

4. Discussion

4.1. Validating the presence of an unusual hydrogen-bond network

We performed the most detailed analysis on structures in space group $P6_3$ since these crystals diffracted to higher resolution and were less radiation-sensitive than those in either space group $P6_5$ or $P2_1$. However, the lower resolution information obtained on other crystal forms is fully consistent with our analysis.

Our structures definitively reveal unusually short hydrogen bonds in the ground state of both wild-type PYP and the isosteric E46Q mutant. These short hydrogen bonds are present in the structures of all eight monomers solved under quite different conditions of temperature, pH, crystallization solvent and intermolecular crystal contacts. They are not a consequence of our crystallographic refinement procedures since all other hydrogen bonds in PYP, including other hydrogen bonds in the chromophore-binding pocket (Table 2), conform to the normal distribution (Fig. 2*a*). Further, these unusually short hydrogen bonds also turn out to be present in certain earlier independently determined structures previously solved at both room and cryogenic temperatures (Borgstahl *et al.*, 1995; Genick *et al.*, 1998; van Aalten *et al.*, 2000). Inspection of the two cryogenic structures (PDB codes 3pyp and 1d7e) shows that they too contain unusually short

hydrogen bonds to the chromophore phenolate oxygen (Table 2). However, the authors did not comment on this key structural observation, perhaps because the original room-temperature PYP structure (PDB code 2phy) did not reveal any short hydrogen bonds.

The 2.5 Å short hydrogen bond from Tyr42 to the chromophore phenolate oxygen is retained in wild-type PYP (strain BN9626), in the E46Q mutant (strain BN9626) and even in the partly disordered wild-type PYP (strain SL-1). Despite the novel dimerization interface in the asymmetric unit of the wild-type PYP (strain SL-1) crystals, the PAS domain core of the protein remains fully intact at pH 9.5 (Fig. 4). The hydrogen bond from Glu46 to the chromophore is shorter in both monomers of the wild-type (strain SL-1) monoclinic crystal than in the wild type at neutral pH in space group $P6_3$ (Table 2). This shorter hydrogen bond from Glu46 is also present in the wild-type $P6_5$ structure at 110 K and pH 6.5 (van Aalten *et al.*, 2000). It therefore appears that while the exceptionally short hydrogen bond from Tyr42 to the chromophore phenolate oxygen retains a length of 2.5 Å under a wide variety of conditions, the length of the hydrogen bond from Glu46 to the chromophore is moderately sensitive to pH, solvent and crystallization conditions.

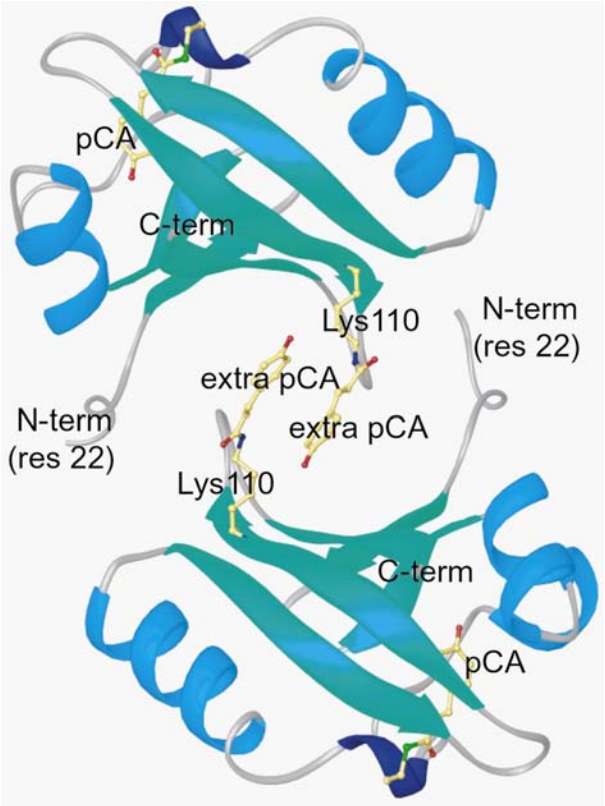

Figure 4
The crystallographic dimer of wild-type PYP (strain SL-1) in space group $P2_1$. The first 21 residues in both monomers are completely disordered. These residues are part of the N-terminal cap (residues 1–28) of PYP that supplements the conserved PAS domain (Taylor & Zhulin, 1999). The asymmetric unit is composed of a dimer stabilized by a hydrophobic interface, which is normally shielded from solvent by the N-terminal cap. This interface is unexpectedly stabilized by an additional coumaric acid chromophore (extra pCA) covalently attached *via* an amide linkage to Lys110. The phenolate oxygen on the additional chromophore forms a hydrogen bond to His108 on the other monomer in the asymmetric unit. The crystals only grew after several months at pH values greater than 9.5. At high pH, the chromophore thioester linkage to Cys69 is prone to hydrolysis (Hoff *et al.*, 1996) and the side chain of lysine becomes uncharged and strongly nucleophilic (Hunter & Ludwig, 1972). Since there was no free chromophore under crystallization conditions, this additional chromophore was obtained by hydrolysis of the thioester linkage between Cys69 and the chromophore of another molecule.

The hydrogen-bond network in PYP serves a structural role that stabilizes the chromophore in its binding pocket. Disruption of this network is critical in the formation of the various photocycle intermediates in the PYP photocycle following the *trans* to *cis* isomerization of the chromophore (Genick, Borgstahl *et al.*, 1997). Once the hydrogen-bond network is broken, the chromophore swings into the solvent and structural changes propagate throughout the entire molecule, with significant changes ultimately occurring at the N-terminal cap and distant regions of the β-sheet (Anderson *et al.*, 2004; Ren *et al.*, 2001; Rubinstenn *et al.*, 1998). Modulating the strength of hydrogen bonds to the chromophore greatly affects the intermediate lifetimes, implying that this network serves as a way for the protein to fine-tune its photocycle and the lifetime of particular intermediates.

4.2. Short hydrogen bonds

Normal hydrogen bonds are characterized by weak electrostatic interactions (Gilli & Gilli, 2000; Hibbert & Emsley, 1990) between hydrogen-bond donor and acceptor atoms separated by ~2.7–3.0 Å (Fig. 2*a*). So-called 'strong' or 'low-barrier' hydrogen bonds are unusually short hydrogen bonds with significant covalent character and an enhanced energy of stabilization (Cleland, 2000; Steiner, 2002), where the energy refers to the difference between the enthalpy of heteroatom interaction with and without the hydrogen bond (Frey, 2001). There is no clear boundary for distinguishing a strong from a normal hydrogen bond, but strong hydrogen bonds are considered to exist when the donor and acceptor atoms are more closely spaced than the sum of their van der Waals radii, which is 2.6 Å for O atoms (Frey, 2001; Hibbert & Emsley, 1990; Steiner, 2002). Although strong hydrogen bonds can exist between any donor–acceptor pair that exhibits very similar pK_a values (Cleland, 2000; Frey, 2001), they are most commonly found between a hydrogen-bond donor and a charged atom, as is the case in PYP. Strong hydrogen bonds have significant thermodynamic stability in isolation, but the energetic stability of a network is greater than the simple sum of the component hydrogen bonds (Steiner, 2002). Cooperative effects such as polarization-enhanced or resonance-assisted hydrogen bonding increase the overall energy of stabilization of the network.

Although there is consensus that hydrogen-bond strength is a continuum with a clear correlation between length and strength given reasonable bond geometry (Gilli & Gilli, 2000; Steiner, 2002), the magnitude of the additional stabilization provided by short hydrogen bonds is debatable. Some strong hydrogen bonds are argued to have exceptional energetic

stability, ~40–80 kJ mol^{-1}, which is utilized during enzymatic catalysis to overcome transition-state barriers and facilitate the formation of covalent bonds (Cleland, 2000; Schowen *et al.*, 2000). Although there is some experimental evidence for this view, numerous other experiments have shown that short hydrogen bonds have only a moderately enhanced stability, of the order of a few kJ mol^{-1} (Perrin & Nielson, 1997; Poi *et al.*, 2003; Usher *et al.*, 1994). Since the energetic stability of individual hydrogen bonds in proteins cannot be measured directly, calculated energies should always be taken with caution (Perrin & Nielson, 1997; Steiner, 2002).

Hydrogen bonds with exceptional energetic stability are very short, but the converse is not always true. Some very short hydrogen bonds have lower stability because of non-canonical bond angles and/or steric clashes between the acceptor and donor atoms (Steiner, 2002). In contrast, both the chromophore and Glx46 in PYP have significant mobility within the context of the protein and can establish an optimal hydrogen-bond distance with good geometry. Relative to the wild type, the chromophore in the E46Q mutant is slightly displaced in order to accommodate the longer hydrogen bond from Gln46 (Figs. 5*c* and 5*d*). Both Tyr42 and Thr50 are displaced together with the chromophore to maintain the wild-type hydrogen-bond lengths; this would not be expected if the exceptionally short hydrogen bond from Tyr42 was too short and destabilizing.

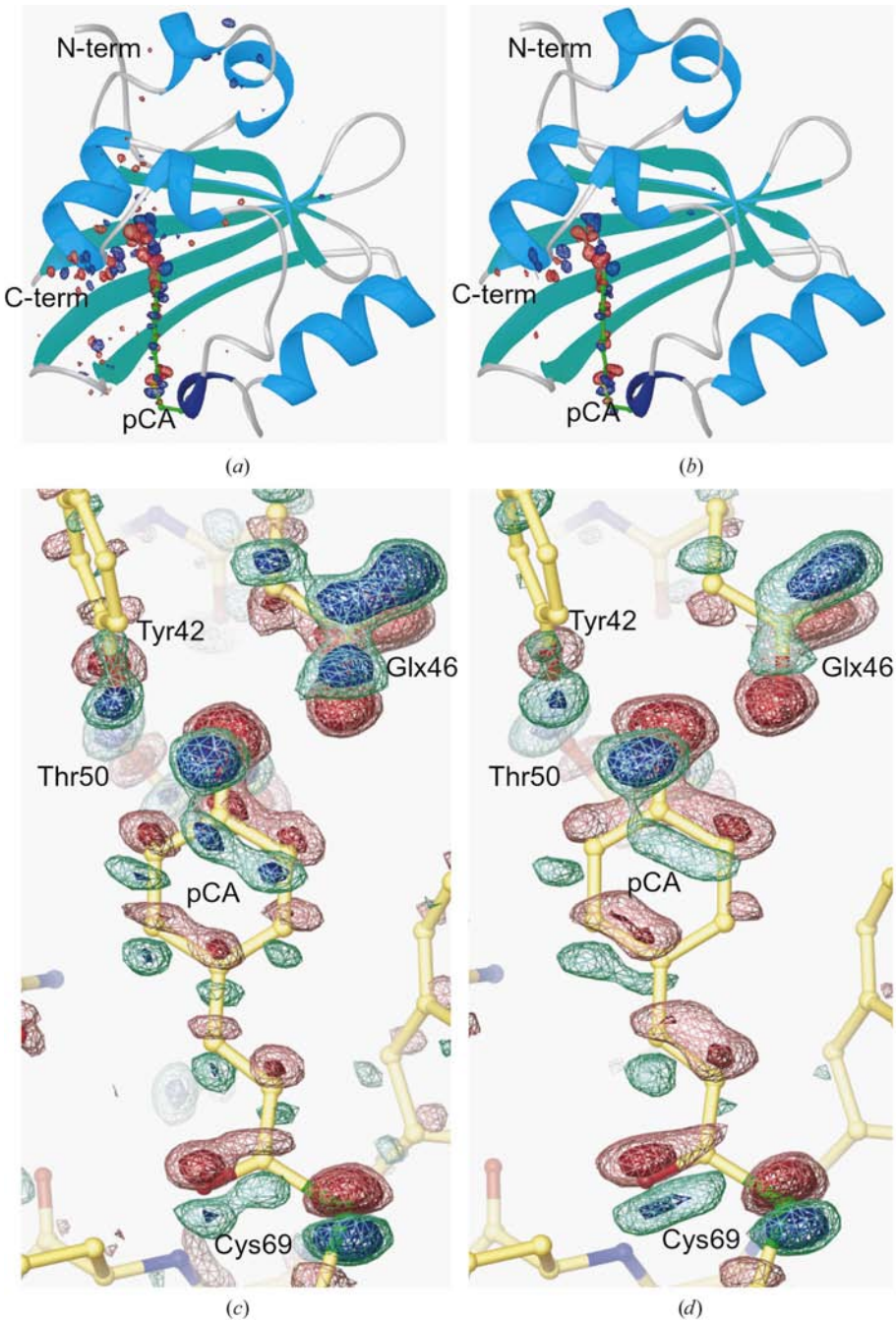

Figure 5
Difference electron-density maps between wild type and E46Q mutant in space group $P6_3$ superimposed on the wild-type structure. Red contours denote negative difference electron density and blue denote positive. The entire molecule contoured at $\pm5\sigma$ is shown at (*a*) 110 K and (*b*) at 295 K. The chromophore-binding pocket contoured at $\pm4\sigma$ and $\pm8\sigma$ is shown at (*c*) 110 K and (*d*) at 295 K.

4.3. Isosteric PYP mutants

High-resolution structures have been determined for three isosteric mutants of PYP (Brudler *et al.*, 2000). The predominant effect of these mutations is to modify the length and presumably the strength of hydrogen bonds to the chromophore while only slightly altering the other aspects of the chromophore-binding pocket. In the E46Q mutant, the hydrogen bond from Gln46 to the phenolate oxygen is of average length (Table 2, Fig. 2), which therefore has less stability than the shorter hydrogen bond from Glu46 in wild type. The hydrogen bond from Tyr42 to the chromophore is unchanged in the E46Q mutant. This implies that the hydrogen bond from Glx46 has little effect on the phenolate anion and its ability to form the unusually short hydrogen bond with Tyr42.

Two earlier high-resolution structures examined the effects of the mutation in the hydrogen-bond network: T50V and Y42F (Brudler *et al.*, 2000). The structure of the T50V mutant was solved to 1.15 Å resolution. The predominant effect of this mutation is loss of the hydrogen bond between Thr50 and Tyr42; the bond length

between Tyr42 and the chromophore phenolate oxygen is 2.62 Å (Brudler *et al.*, 2000). The authors compared this length with that of the corresponding hydrogen bond in the original wild-type room-temperature structure of 2.71 Å (Borgstahl *et al.*, 1995) and therefore commented that this hydrogen bond is shorter and slightly stronger in the T50V mutant. On the contrary, our analysis at higher crystallographic resolution shows that the hydrogen bond from Tyr42 to the chromophore in the T50V mutant is significantly longer than in the wild type, 2.62 Å compared with 2.50 Å (Table 2). This demonstrates how the hydrogen bonds to the chromophore are a finely tuned network in which the hydrogen bond from Thr50 to Tyr42 in wild-type PYP is an important factor in stabilizing the short hydrogen bond from Tyr42 to the chromophore phenolate oxygen.

The structure of the Y42F mutant was solved to 1.10 Å resolution (Brudler *et al.*, 2000). Although the Y42F mutation is nearly isosteric, Phe42 cannot form hydrogen bonds to the chromophore phenolate oxygen or to Thr50. The Y42F mutant acquires a new hydrogen bond from Thr50 to the chromophore that is of normal length, 2.79 Å. The ground state of the Y42F mutant in solution is in equilibrium with a relatively disordered species that has a protonated *trans* chromophore (Brudler *et al.*, 2000; Imamoto *et al.*, 2001). This disordered species is akin to pB^{dark}, a ground-state conformation of PYP similar to the pB photocycle intermediate formed at low pH without absorption of a photon (Hoff *et al.*, 1997). Formation of this disordered state in the Y42F mutant implies that the hydrogen bond from Tyr50 to the chromophore plays an important role in securing the chromophore in its binding pocket, preventing the protein from establishing its active conformation unless it is photoactivated.

The E46Q, T50V and Y42F isosteric mutants all either completely remove a hydrogen bond or lengthen (and by hypothesis, weaken) a bond in the hydrogen-bond network to the chromophore. Since the network is completely broken in the pB intermediate (Genick, Borgstahl *et al.*, 1997), the mutations would be expected to lower the energetic barrier for transition to the pB intermediate. Indeed, the rate of pB intermediate formation is increased approximately fivefold over that in wild type in the E46Q mutant and approximately threefold in the T50V mutant (Genick, Devanathan *et al.*, 1997). Formation of the pB intermediate in the Y42F mutant occurred on a microsecond timescale and was therefore too fast to accurately measure at room temperature and neutral pH (Brudler *et al.*, 2000). Although much emphasis has been placed on the key role of Glu46 in the structure and photocycle of PYP (Imamoto *et al.*, 1997; Xie *et al.*, 1996;, 2001), it is not surprising that modifying Tyr42 has a greater effect on the photocycle given its unusually short hydrogen bond to the chromophore.

5. Conclusions

Our analysis of eight crystallographic PYP monomers has clearly identified unusual hydrogen bonds to the chromophore phenolate oxygen. These short hydrogen bonds are key structural elements that control photocycle kinetics and hence the signaling efficacy of this PAS-domain blue-light photoreceptor. The hydrogen bonds are stabilized by the finely tuned electronic environment in the ground-state chromophore-binding pocket. Disrupting this environment *via* isosteric mutations weakens these bonds and hastens the rate at which they break during the PYP photocycle. These measurements enhance our understanding of the PYP photocycle mechanism and document the influence of environment on short hydrogen bonds in biological macromolecules.

We thank Carl Correll, Anthony Kossiakoff, Marvin Makinen and Wouter Hoff for their insights and many productive discussions. Thanks to the staff at BioCARS and fellow graduate students Sudarshan Rajagopal and Jason Key for help with data collection and advice. This work was supported by National Institutes of Health Grant GM36452 (KM); the BioCARS facility at the Advanced Photon Source is supported by National Institutes of Health Grant RR07707 (KM).

References

Aalten, D. M. F. van, Crielaard, W., Hellingwerf, K. J. & Joshua-Tor, L. (2000). *Protein Sci.* **9**, 64–72.
Anderson, S., Srajer, V., Pahl, R., Rajagopal, S., Schotte, F., Anfinrud, P., Wulff, M. & Moffat, K. (2004). Submitted.
Borgstahl, G. E., Williams, D. R. & Getzoff, E. D. (1995). *Biochemistry*, **34**, 6278–6287.
Brudler, R., Meyer, T. E., Genick, U. K., Devanathan, S., Woo, T. T., Millar, D. P., Gerwert, K., Cusanovich, M. A., Tollin, G. & Getzoff, E. D. (2000). *Biochemistry*, **39**, 13478–13486.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cleland, W. W. (2000). *Arch. Biochem. Biophys.* **382**, 1–5.
Cusanovich, M. A. & Meyer, T. E. (2003). *Biochemistry*, **42**, 965–970.
Devanathan, S., Lin, S., Cusanovich, M. A., Woodbury, N. & Tollin, G. (2000). *Biophys. J.* **79**, 2132–2137.
Devanathan, S., Pacheco, A., Ujj, L., Cusanovich, M., Tollin, G., Lin, S. & Woodbury, N. (1999). *Biophys. J.* **77**, 1017–1023.
Dux, P., Rubinstenn, G., Vuister, G. W., Boelens, R., Mulder, F. A., Hard, K., Hoff, W. D., Kroon, A. R., Crielaard, W., Hellingwerf, K. J. & Kaptein, R. (1998). *Biochemistry*, **37**, 12689–12699.
Frey, P. A. (2001). *Magn. Reson. Chem.* **39**, S190–S198.
Genick, U. K., Borgstahl, G. E., Ng, K., Ren, Z., Pradervand, C., Burke, P. M., Srajer, V., Teng, T.-Y., Schildkamp, W., McRee, D. E., Moffat, K. & Getzoff, E. D. (1997). *Science*, **275**, 1471–1475.
Genick, U. K., Devanathan, S., Meyer, T. E., Canestrelli, I. L., Williams, E., Cusanovich, M. A., Tollin, G. & Getzoff, E. D. (1997). *Biochemistry*, **36**, 8–14.
Genick, U. K., Soltis, P. S., Kuhn, M., Canestrelli, I. L. & Getzoff, E. D. (1998). *Nature (London)*, **392**, 206–209.
Gilli, G. & Gilli, P. (2000). *J. Mol. Struct.* **552**, 1–15.
Harris, T. K. & Turner, G. J. (2002). *IUBMB Life*, **53**, 85–98.
Hellingwerf, K. J., Hendriks, J. & Gensch, T. (2003). *J. Phys. Chem. A*, **107**, 1082–1094.
Hibbert, F. & Emsley, J. (1990). *Adv. Phys. Org. Chem.* **26**, 255.

Hoff, W. D., Devreese, B., Fokkens, R., Nugteren-Roodzant, I. M., van Beeumen, J., Nibbering, N. & Hellingwerf, K. J. (1996). *Biochemistry*, **35**, 1274–1281.
Hoff, W. D., van Stokkum, I. H. M., Gural, J. & Hellingwerf, K. J. (1997). *Biochim. Biophys. Acta*, **1322**, 151–162.
Hunter, M. J. & Ludwig, M. L. (1972). *Methods Enzymol.* **25**, 14671–14678.
Imamoto, Y., Koshimizu, H., Mihara, K., Hisatomi, O., Mizukami, T., Tsujimoto, K., Kataoka, M. & Tokunaga, F. (2001). *Biochemistry*, **40**, 4679–4685.
Imamoto, Y., Mihara, K., Hisatomi, O., Kataoka, M., Tokunaga, F., Bojkova, N. & Yoshihara, K. (1997). *J. Biol. Chem.* **272**, 12905–12908.
Kim, M., Mathies, R. A., Hoff, W. D. & Hellingwerf, K. J. (1995). *Biochemistry*, **34**, 12669–12672.
Kort, R., Hoff, W. D., van West, M., Kroon, A. R., Hoffer, S. M., Vlieg, K. H., Crielaand, W., van Beeumen, J. J. & Hellingwerf, K. J. (1996). *EMBO J.* **15**, 3209–3218.
Kroon, A. R., Hoff, W. D., Fennema, H. P., Gijzen, J., Koomen, G. J., Verhoeven, J. W., Crielaard, W. & Hellingwerf, K. J. (1996). *J. Biol. Chem.* **271**, 31949–31956.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
McRee, D. E., Tainer, J. A., Meyer, T. E., van Beeumen, J., Cusanovich, M. A. & Getzoff, E. D. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 6533–6537.
Meyer, T. E. (1985). *Biochim. Biophys. Acta*, **806**, 175–183.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrin, C. L. & Nielson, J. B. (1997). *Annu. Rev. Phys. Chem.* **48**, 511–544.
Poi, M. J., Tomaszewski, J. W., Yuan, C. H., Dunlap, C. A., Andersen, N. H., Gelb, M. H. & Tsai, M. D. (2003). *J. Mol. Biol.* **329**, 997–1009.
Ren, Z., Perman, B., Srajer, V., Teng, T.-Y., Pradervand, C., Bourgeois, D., Schotte, F., Ursby, T., Kort, R., Wulff, M. & Moffat, K. (2001). *Biochemistry*, **40**, 13788–13801.
Rubinstenn, G., Vuister, G. W., Mulder, F. A. A., Dux, P. E., Boelens, R., Hellingwerf, K. J. & Kaptein, R. (1998). *Nature Struct. Biol.* **5**, 568–570.
Schowen, K. B., Limbach, H. H., Denisov, G. S. & Schowen, R. L. (2000). *Biochim. Biophys. Acta*, **1458**, 43–62.
Sheldrick, G. (1997). *SHELXL*97. University of Göttingen, Germany.
Sprenger, W. W., Hoff, W. D., Armitage, J. P. & Hellingwerf, K. J. (1993). *J. Bacteriol.* **175**, 3096–3104.
Steiner, T. (2002). *Angew. Chem. Int. Ed.* **41**, 48–76.
Taylor, B. L. & Zhulin, I. B. (1999). *Microbiol. Mol. Biol. Rev.* **63**, 479–506.
Usher, K. C., Remington, S. J., Martin, D. P. & Drueckhammer, D. G. (1994). *Biochemistry*, **33**, 7753–7759.
Xie, A., Hoff, W. D., Kroon, A. R. & Hellingwerf, K. J. (1996). *Biochemistry*, **35**, 14671–14678.
Xie, A., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J. & Hoff, W. D. (2001). *Biochemistry*, **40**, 1510–1507.
Zhou, Y., Ujj, L., Meyer, T. E., Cusanovich, M. A. & Atkinson, G. H. (2001). *J. Phys. Chem. A*, **105**, 5719–5726.