Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and 2.2 Å resolution structure of R-phycoerythrin from *Gracilaria chilensis*: a case of perfect hemihedral twinning

C. Contreras-Martel,[a,b] J. Martinez-Oyanedel,[a] M. Bunster,[a]* P. Legrand,[b] C. Piras,[b] X. Vernede[b] and J. C. Fontecilla-Camps[b]

[a]Laboratorio de Biofísica Molecular, Departamento de Biología Molecular, Facultad de Ciencias Biológicas, Universidad de Concepción, Chile, and [b]Laboratoire de Cristallographie et Cristallogenèse des Protéines, Institut de Biologie Structurale J. P. Ebel, CEA/CNRS, Grenoble, France

Correspondence e-mail: mbunster@udec.cl

Received 15 May 2000
Accepted 25 October 2000

PDB Reference: R-phycoerythrin, 1eyx.

R-phycoerythrin, a light-harvesting component from the red algae *Gracilaria chilensis*, was crystallized by vapour diffusion using ammonium sulfate as precipitant agent. Red crystals grew after one week at 293 K and diffracted to 2.70 Å resolution. Three serial macroseeding assays were necessary to grow a second larger crystal to dimensions of 0.68 × 0.16 × 0.16 mm. This crystal diffracted to 2.24 Å resolution using synchrotron radiation at beamline BM14 of the European Synchrotron Radiation Facility (ESRF) at Grenoble, France and was used for structure determination. Data were collected at 100 K to a completeness of 98.6%. The crystal was trigonal, space group $R3$, with unit-cell parameters $a = b = 187.3$, $c = 59.1$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Data treatment using the *CCP*4 suite of programs indicated that the crystal was twinned ($\langle I^2\rangle/\langle I\rangle^2 = 1.41$). Molecular replacement was performed with *AMoRe* using the R-phycoerythrin from *Polysiphonia urceolata* [Chang *et al.* (1996), *J. Mol. Biol.* **249**, 424–440] as a search model. In order to overcome the twinning problem, *SHELX*97 was used for the crystallographic refinement. The twin fraction was 0.48, indicating a nearly perfect hemihedrally twinned crystal. The final R_{work} and R_{free} factors are 0.16 and 0.25, respectively. All the residues and chromophores of the α- and β-chains are well defined in the electron-density maps. Some residues belonging to the γ-linker are also recognizable.

1. Introduction

Phycobiliproteins (PBPs) are brilliantly coloured highly fluorescent proteins formed by an apoprotein bound to a linear tetrapyrrolic prosthetic chromophore (Glazer, 1989). The PBPs are present in cyanobacteria, blue–green and red algae, where they are organized in a macromolecular complex, the phycobilisome (PBS; Gantt, 1981). Three main proteins compose the PBS: phycoerythrin (PE), phycocyanin (PC) and allophycocyanin (APC). In addition, several colourless linker proteins are found in this macromolecular array. The main function of PBS is light harvesting and energy transfer to the photoreaction centres (Kursar *et al.*, 1983). Each PBP may contain one or more of four specific chromophores called phycoerythrobilin (PEB), phycocyanobilin (PCB), phycobiliviolin (PXB) and phycourobilin (PUB), covalently linked to cysteine residues through a thioether bond. The various PBPs are highly homologous and probably originated from a common ancestor (Apt *et al.*, 1995).

The PBS of *G. chilensis*, a red eukaryotic algae, contains the three proteins R-phycoerythrin (R-PE), R-phycocyanin (R-PC) and R-allophycocyanin (R-APC) (Bunster *et al.*, 1997). In the blue–green algae *Porphyridium sordidum,*

B-phycoerythrin (B-PE) is made up of three $\alpha\beta$ heterodimers plus a γ-subunit that has been considered as a linker protein (L_R) responsible for the packing of the phycobilisome (Ficner & Huber, 1993). B-PE has two PEBs per α-chain and two PEBs and a PUB per β-chain (Glazer, 1989). The structures of B-PE from *Porphyridium sordidum* and *Porphyridium cruentum* have been solved to 2.3 Å (Ficner *et al.*, 1992) and 2.2 Å resolution (Ficner & Huber, 1993), respectively. At the beginning of the present project there were no R-PE structures reported. There are now two PE structures deposited in the Protein Data Bank: R-PE from *Polysiphonia urceolata* at 2.8 Å resolution (PDB code 1lia) solved by Chang *et al.* (1996) and R-PE from *Griffithsia monilis* at 1.9 Å resolution (PDB code 1b8d) solved by Ritter *et al.* (1999). For a review on PBPs, see Glazer (1989) and Betz (1997).

Here, we report the crystallization, molecular-replacement structure solution and crystallographic refinement of *Gracilaria chilensis* R-PE from a merohedrally twinned crystal (Koch, 1992; Redinbo & Yeates, 1993). There are few reported examples of structures solved from twinned crystals in macromolecular crystallography and all of them correspond to hemihedral twins (Yeates, 1997; Yeates & Fam, 1999). Molecular replacement is a good approach to solve a structure in the case of a 'perfect twin' (domain ratio $\simeq 0.5$; Valegård *et al.*, 1998; Chandra *et al.*, 1999).

```
Alpha Chains
      1               15 16              30 31              45 46              60 61              75
polbo MKSVITTTISAADAA GRYPSTSDLQSVQGN IQRAAARLEAAEKLG SNHEAVVKEAGDACF SKYGYLKNPGEAGEN
seq   MKSVITTVISAADSA GRFPSSSDLESVQGN IQRASARL?AA??LA ?N
den   MKSVITTVISAADSA GRFPSSSDLESVQGN IQRASARLEAAEKLA SNHEAVVKEAGDACF GKYGYLKNPGEAGEN
grimo MKSVITTTISAADAA GRFPSSSDLESIQGN IQRAAARLEAAQKLS GNHEAVVKEAGDACF AKYSYLKNAGEAGDS
porso MKSVITTVISAADAA GRFPSNSDLESIQGN IQRSAARLEAAEKLA GNHEAVVKEAGDACF AKYAYLKNPGEAGEN

      76          90 91              105 106             120 121             135 136             150
polbo QEKINKCYRDIDHYM RLINYTLVVGGDGPL DEWGIAGAREVYRTL NLPSAAYIAAFVFTR DRLCIPRDMSAQAGV
seq                   LVNYSLVIGGTGPL DEWGIAGAREVYRTL NLPTSA                    MSAQAGV
den   QEKINKCYRDIDHYM RLVNYSLVIGGTGPL DEWGIAGAREVYRTL NLPTSAYIAAFAFTR DRLCGPRDMSAQAGV
grimo PEKINKCYRDIDHYM RLINYSLVVGGTGPV DEWGIAGSREVYRAL NLPGSAYIAAFTFTR DRLCVPRDMSSQAGV
porso QEKINKCYRDVDHYM RLVNYDLVVGGTGPL DEWGIAGAREVYRTL NLPTSAYVASIAYTR DRLCVPRDMSAQAGV

      151       164
polbo EFCTALDYLINSLS
seq   EY?T?LDYIINSLS
den   EYSTALDYIINSLS
grimo EFTSALDYVINSLC
porso EFSAYLDYLINALS

Beta Chains
      1               15 16              30 31              45 46              60 61              75
polbo MLDAFSRVVVNSDSK AAYVSGSDLQALKTF INDGNKRLDAVNYIV SNSSCIVSDAISGMI CENPGLITPGGNCYT
seq   MLDAFSRV??NAD?? ??YVGG?DL???R?? IS????RL?A
den   MLDAFSRVISNADAK AAYVGGSDLQALRTF ISDGNKRLDAVNYIV SNSSCIVSDAISGMI CENPGLITPGGNCYT
grimo MLDAFSRVVVTSDAK AAYVGGSDLQSLKSF INDGNKRLDAVNYIV SNASCIVSDAVSGMI CENPGLIAPGGXCYT
porso MLDAFSRVVVNSDAK AAYVGGSDLQALKSF IADGNKRLDSVNAIV SNASCMVSDAVSGMI CENPGLISPGGNCYT

      76          90 91              105 106             120 121             135 136             150
polbo NRRMAACLRDGEIIL RYVSYALLAGDASVL EDRCLNGLKETYIAL GVPTNSTVRAVSIMK AAAVCFISNTASQRK
seq                                                                        VGAFISNTASQRK
den   NRRMAACLRDGEIIL RYISYALLAGDSSVL EDRCLNGLKETYIAL GVPTNSTVRAVSIMK AAVGAFISNTASQRK
grimo NRRMAACLRDGEIIL RYVSYALLAGDSSVL DDRCLNGLKETYIAL GVPTASSSRAVSIMK ATATAFITNTASGRK
porso NRRMAACLRDGEIIL -YVSYALLAGDASVL EDRCLNGLKETYIAL GVPTNSSIRAVSIMK AQAVAFITNTATERK

      151       165 166       177
polbo VEVIEGDCSALASEV ASYCDRVVAAVS
seq
den   GEVIEGDCSALAAEI ASYCDRISAAVS
grimo VEVAAGDCQALQAEA ASYFDKVGSSID
porso MSFAAGDCTSLASEV ASYFDRVGAAIS
```

Figure 1
Determined amino-acid sequences of α- and β-chains of R-PE from *G. chilensis*. The partial chemical sequence (seq) and the sequence based on electron-density maps (den) are shown. They are compared with homologous sequences of *Polysiphonia boldii* (polbo, model sequence); *G. monilis* (grimo) and *Porphyridium sordidum* (porso) by a multiple-sequence alignment (PIMA1.4) (Smith & Smith, 1992). The conserved changes are shown in blue and the non-conserved changes in red.

2. Materials and methods

2.1. Protein purification

The proteins were extracted from fresh algae (Gantt *et al.*, 1979; Bunster *et al.*, 1997). Three coloured PBPs were resolved by anion-exchange chromatography (TSK-GEL DEAE-5PW 15 × 21.5 mm ID, Tosohaas) with a HPLC system (BIO LC, Dionex). The column was equilibrated with 5 m*M* potassium phosphate buffer pH 7.0 and eluted with a linear gradient from 0 to 1.0 *M* sodium chloride. The protein concentration was determined by the BCA method (Micro BCA Protein Assay, Pierce; Smith *et al.*, 1985). The purity and homogeneity of the fractions was analyzed by SDS and native PAGE (PhastGel, Homogeneous 20/2–150 kDa, Pharmacia).

2.2. Crystallization

Crystals were grown by the vapour-diffusion method using sitting-drop insets (Linbro plates, DROP, SA). 1 ml of 16–18 % saturation ammonium sulfate solution in 200 m*M* HEPES buffer pH 7.0, 500 m*M* sodium chloride, 50 m*M* potassium chloride and 15 m*M* sodium azide was used in the reservoir. The drops were formed by mixing 12–15 mg ml^{-1} protein with reservoir solution in either 1:1 or 2:1 ratios. Red crystals appeared after 3–4 d and continued to grow for three weeks at 294 K as hexagonal rods to dimensions of 0.30 × 0.04 × 0.04 mm. After three serial macroseeding assays spanning a period of three weeks, crystals grew to maximum dimensions of 0.68 × 0.16 × 0.16 mm.

2.3. Data collection and space-group assignment

The crystals were soaked in a cryoprotective solution containing 40%(*v*/*v*) glycerol in the crystallization mother liquor and cooled to 100 K. A data set was collected from a crystal of dimensions 0.4 × 0.1 × 0.1 mm to 2.7 Å resolution at the ID9 beamline of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) using a CCD detector. Data from a second larger crystal was collected to 2.2 Å resolution at BM14 of the ESRF using a

MAR345 imaging-plate detector. These data were indexed using *XDS*98 (Kabsch, 1993) and *HKL*2000 (Otwinowski & Minor, 1997) and were reduced using the *CCP*4 set of programs (Collaborative Computational Project, Number 4, 1994) and *SCALEPACK* (Otwinowski & Minor, 1997). The space group *R*3 was assigned for both crystals directly by *XDS*98 and *HKL*2000. To check this space-group assignment, the *hk*0 and *hk*3 planes were generated using *HKLVIEW* (Collaborative Computational Project, Number 4, 1994).

2.4. Statistical analysis and twin detection

Both data sets were analyzed using *CCP*4 programs (Collaborative Computational Project, Number 4, 1994). The Stanley factor $\langle I^2\rangle/\langle I\rangle^2$ (Stanley, 1972; Yeates, 1997) and Wilson's diffraction distribution (Wilson, 1949) were obtained from *TRUNCATE* (Collaborative Computational Project, Number 4, 1994).

2.5. Molecular replacement and refinement

Molecular replacement was carried out with *AMoRe* (Navaza, 1994) using the R-PE $(\alpha\beta)_2$ atomic coordinates of Chang *et al.* (1996) (PDB code 1lia). As the amino-acid sequence of *G. chilensis* R-PE was not available at this time, the sequence of the R-PE from *Polysiphonia boldii* (Roell & Morse, 1993) was initially used. This was considered a reasonable option given the high (>80%) homology generally observed between phycoerythrins (Apt *et al.*, 1995) (Fig. 1).

*SHELX*97 (Sheldrick & Schneider, 1997) was used for the crystallographic refinement (TWIN and BASF options). To calculate $R_{\rm free}$, the test reflections were selected in thin resolution shells; to avoid possible correlations introduced by the twinning, all the twin-related pairs belonged to either the reference or the working set (Sheldrick & Schneider, 1997). Five conjugate-gradient cycles were performed in each step of the refinement. Chromophore-definition restraints and water-molecule assignments were performed using *SHELX*97 (*SHELXPRO* and *SWAT*, respectively). Omit maps were systematically used during model building in order to minimize the strong model bias of phases calculated from the molecular-replacement solution (McRee, 1999*a*). The omit maps were calculated after *SHELX*97 refinement and examined with *TURBO-FRODO* (Roussel & Cambillau, 1989) and *XtalView*/*Xfit* v4.0.7 (McRee, 1999*b*) on a graphics workstation. At advanced levels of the refinement, non-crystallographic symmetry restraints (NCSY option) were introduced between homologous chains (Sheldrick & Schneider, 1997). The stereochemical parameters were checked with *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT IF* (Vriend, 1990).

2.6. Chain separation and sequencing

PE was denatured in 4 *M* urea, 0.4%(*v*/*v*) acetic acid, 2 m*M* β-mercaptoethanol, 5 m*M* potassium phosphate buffer pH 1.90 (Troxler *et al.*, 1981). The α- and β-chain were separated using a 4–10 *M* urea gradient in a HPLC (BIO LC, Dionex), Mono-S HR 5/5 (Pharmacia) column. The process was monitored at 566 nm absorbance. Partial amino-acid sequence information (J. Gagnon, personal communication) was obtained by automated Edman degradation.

3. Results and discussion

3.1. Crystallization

Initial crystallization trials resulted in extensive nucleation and only small crystals were obtained. This problem was

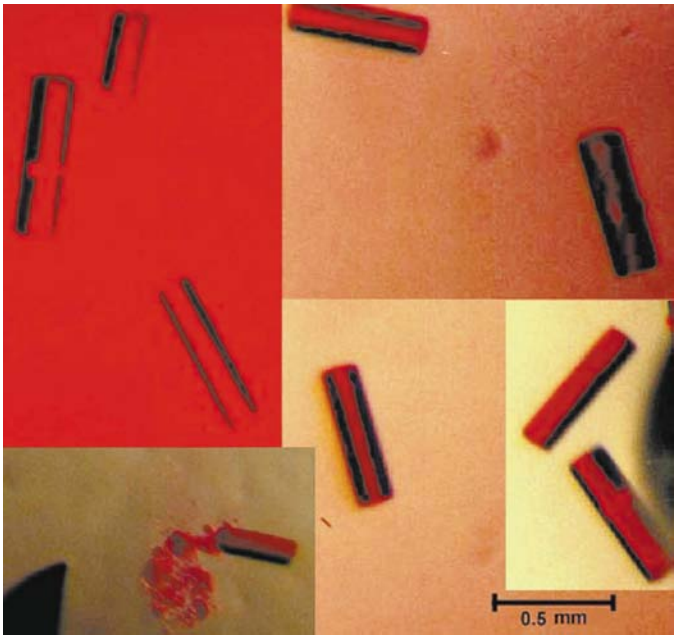

Figure 2
Hexagonal phycoerythrin crystals. Different levels of macroseeding are shown.

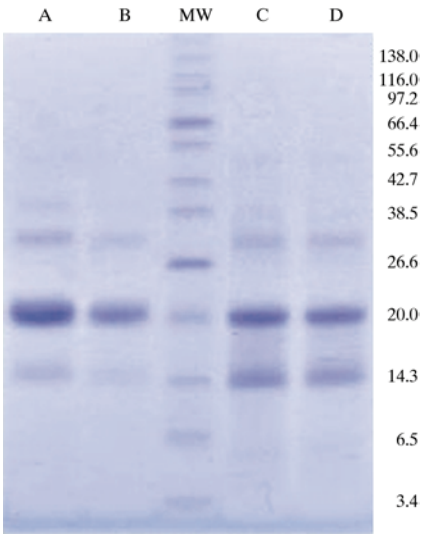

Figure 3
SDS–PAGE of R-phycoerythrin. Lanes *A* and *B*, resuspended protein from crystals; lanes *C* and *D*, fresh protein after DEAE-HPLC purification. Lane MW: molecular-weight markers (values in kDa are indicated on the right side).

solved through macroseeding and the addition of sodium and potassium chloride. The best results were obtained by three successive macroseeding steps in the presence of 0.4–0.5 *M* NaCl (Fig. 2). Under these conditions, nucleation was reduced by about 50%. On the other hand, the use of potassium chloride improved the overall aspect of the crystals.

The homogeneity of the protein sample, from both the original solution and dissolved crystals, was verified by gel electrophoresis. The molecular weights of the bands observed using SDS–PAGE (Fig. 3) were those expected for α- and β-chains (17–18 kDa) and for the γ-chain (30 kDa) (Glazer, 1985; Ficner & Huber, 1993, Ducret *et al.*, 1996). However, there is an unexpected band of about 14 kDa observed in both the original protein solution and dissolved crystals that could be a previously unreported R-PE protein linker (L_R). Two small linkers of 7 and 10 kDa have been described for APC of *Synechocystis* sp. PCC6714 which may correspond to L_C and L_R, respectively (Ajlani & Vernotte, 1998). Similarly, 8.9 kDa L_C and L_R polypeptides have been reported in the case of the cyanobacterium phycobilisome (Ducret *et al.*, 1996). Further characterization of the 14 kDa band is under way.

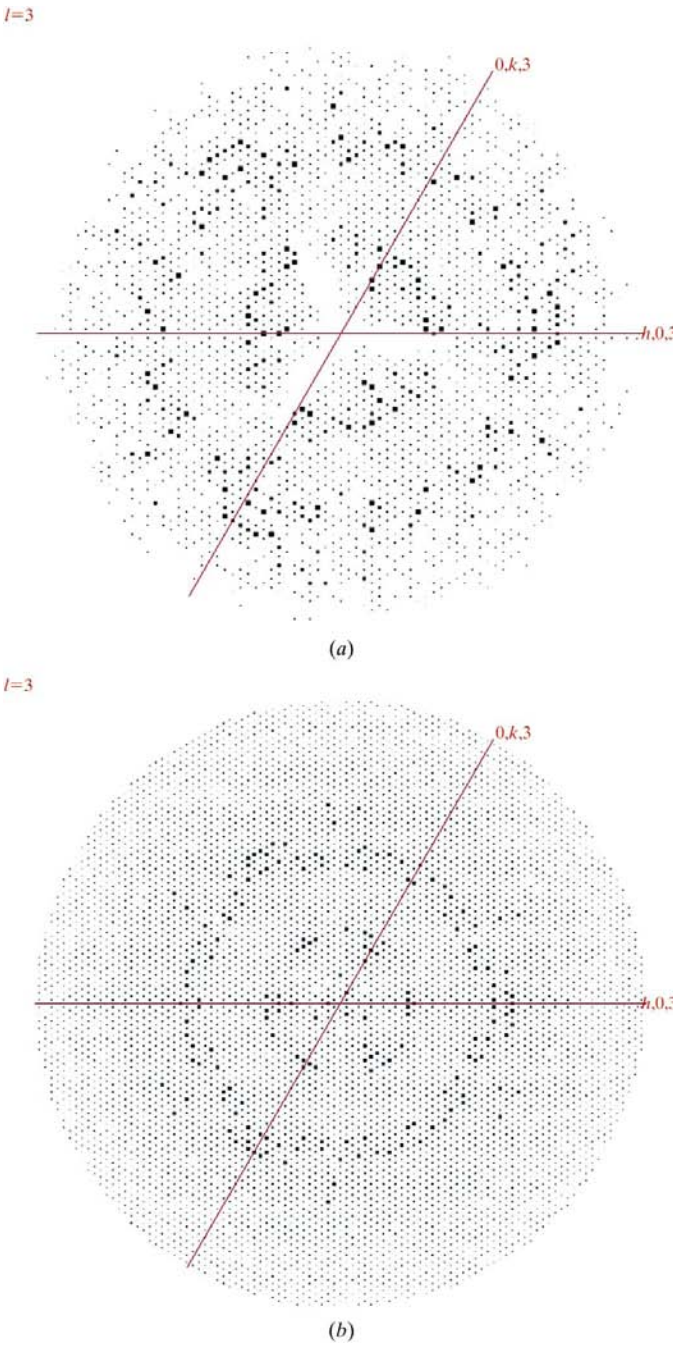

Figure 4
Patterns of data to (*a*) 2.7 Å (untwinned) and (*b*) 2.2 Å (twinned) resolution from *HKLVIEW* (Collaborative Computational Project, Number 4, 1994) (*hk*3 planes).

Table 1
Data-collection statistics and crystallographic refinement statistics.

Values in parentheses are for the last shell, 2.28–2.16 Å.

Unit-cell parameters (Å,°)	$a = b = 187.3$, $c = 59.1$, $\alpha = \beta = 90$, $\gamma = 120$
Space group	*R*3
Resolution (Å)	2.25
Wavelength (λ) (Å)	0.9783
No. of recorded observations	123143
No. of unique reflections	38751
Completeness, 42.55–2.16 Å (%)	98.7 (89.4)
R_{sym}† (%)	10.1 (30.9)
$I/\sigma(I)$	7.3 (2.8)
Multiplicity	3.1 (2.0)
Stanley distribution ($\langle I^2\rangle/\langle I\rangle^2$)	1.41
Twin fraction	0.48
Matthews number (Å^3 Da^{-1})	2.52
R_{work}‡, all data	0.1801
R_{work}‡, $F > 4\sigma(F)$ data	0.1625
R_{free}§(5%), all data	0.2796
R_{free}§ (5%), $F > 4\sigma(F)$ data	0.2557
No. of protein atoms	5151
No. of chromophore atoms	448
No. of solvent molecules	110
No. of sulfate molecules	2
RMS deviation, bond lengths (Å)	0.005
RMS deviation, bond angles (°)	1.062
RMS deviation, side-chain planarity (Å)	0.003
Average *B* factor, main chain (Å^2)	21.602
Average *B* factor, side chain (Å^2)	22.515

† $R_{sym} = 100\sum_{hkl}\sum_i|I_i(hkl) - \bar{I}(hkl)|/\sum_{hkl}\sum_i I_i(hkl)$ for *i* observations of a given reflection, where $\bar{I}(hkl)$ is the average intensity of *i*. ‡ $R_{work} = 100\sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes. § $R_{free} = 100\sum_{hkl\in T}||F_{obs}| - |F_{calc}||/\sum_{hkl\in T}|F_{obs}|$ for a *hkl* reflection belonging to a test set *T* of unique reflections.

3.2. Data analysis

As indicated above, the indexing of the 2.7 and 2.2 Å resolution data indicated that the crystals belong to the *R*3 space group. Statistical analysis using *TRUNCATE* (Collaborative Computational Project, Number 4, 1994) showed $\langle I^2\rangle/\langle I\rangle^2$ ratios of 1.41 and 3.18 for the 2.2 and 2.7 Å resolution data sets, respectively, suggesting the presence of hemihedral twinning in the former; it has been determined that the $\langle I^2\rangle/\langle I\rangle^2$ ratio should be about 2.0 for untwinned data and 1.5 for twinned data (Stanley, 1972; Yeates, 1997). In hemihedral twinned crystals there are two parallel domains, h_1 and h_2, related by an extra symmetry operation that makes

both domains contribute to the diffraction pattern by $I_{obs} = \alpha I(h_1) + (1 - \alpha)I(h_2)$, where α corresponds to the domain-fraction contribution or 'twin fraction'. Until recently, twinning of protein crystals had seldom been reported (Redinbo & Yeates, 1993). Usually, the twinned crystals were discarded or the twin went undetected and the structure was solved with some difficulty (Luecke *et al.*, 1998; Ibuka *et al.*, 1999; Esposito *et al.*, 1999; Yusupov *et al.*, 1999). In the last few years, however, there have been several reports on structures solved from twinned protein crystals (Gomis-Rüth *et al.*, 1995; Ito *et al.*, 1995; Valegård *et al.*, 1998; Frazão *et al.*, 1999; Yeates & Fam, 1999; Breyer *et al.*, 1999; Dumas *et al.*, 1999; Chandra *et al.*, 1999; Yang *et al.*, 2000).

*R*3 is a fairly common true space group for twinned crystals (Sweet *et al.*, 1977; Ficner *et al.*, 1992; Ritter, 1999; Valegård *et al.*, 1998). A perfect hemihedral twinned crystal in the *R*3

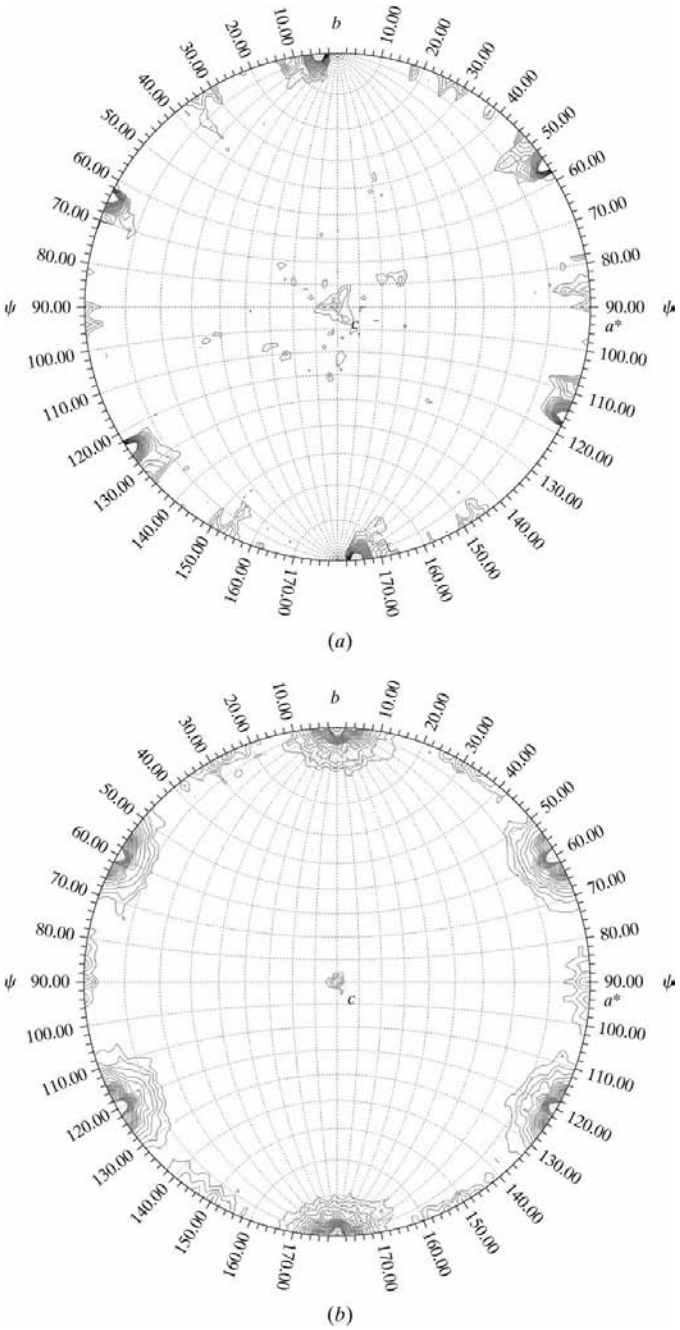

Figure 5
Patterson self-rotation functions of (*a*) 2.7 Å (untwinned) and (*b*) 2.2 Å resolution (twinned) diffraction data calculated with an integration radius of 20 Å and data between 10.0 and 3.5 Å resolution (*GLRF*; Tong & Rossmann, 1990).

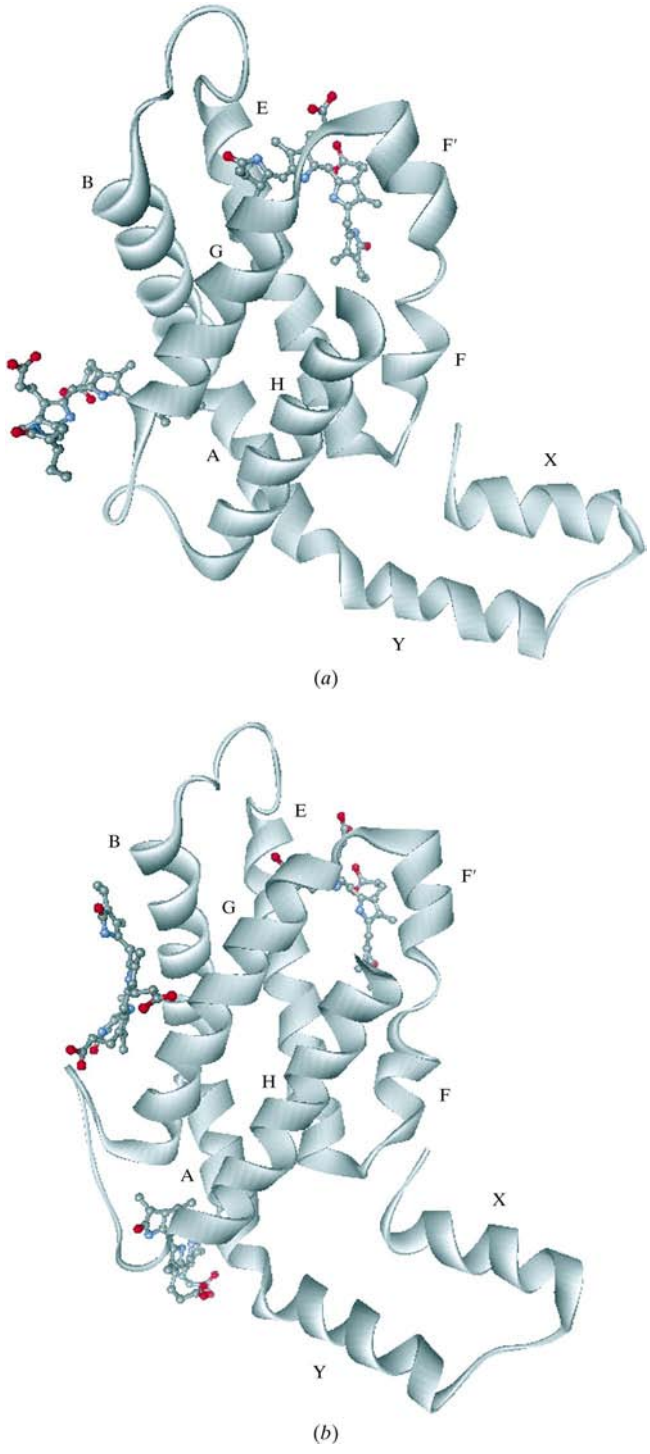

Figure 6
Ribbon representation of chains α (*a*) and β (*b*) of R-phycoerythrin from *G. chilensis*. The code of helical segments was assigned according to Schirmer *et al.* (1986) and elsewhere. Chromophores are depicted as balls and sticks. Figure prepared with the program *WebLab*3.5 MSI.

space group can be erroneously indexed as $R32$, as the only possible twin symmetry operator for $R3$, k, h, $-l$, generates a twofold axis on the diagonal between the a and b axes (Chandra *et al.*, 1999). Figs. 4(*a*) and 4(*b*) depict a comparison of $hk3$ planes for our untwinned and twinned data. The pattern in Fig. 4(*a*) shows the true threefold symmetry along the l axis ($R3$, trigonal indexing). For twinned data, as shown in Fig. 4(*b*), the twin operation generates an approximate sixfold axis that underlies the threefold symmetry. There is, however, no true twofold axis in the diagonal between the a and b axes as required in $R32$. In addition, the Matthews number (V_M; Matthews, 1968) of 2.52 Å^3 Da^{-1} also indicates that the $R3$ space group assignment is correct (Table 1); V_M values below 1.68 Å^3 Da^{-1} are found when the assigned symmetry is higher than the correct symmetry (Redinbo & Yeates, 1993).

Examination of the untwinned and twinned data using the graphic representation of Tong & Rossman (1990) for $\kappa = 180^\circ$ with the program *GLRF* (Tong & Rossman, 1990) indicates, as expected, the presence of twofold axes perpendicular to the unique axis and separated by 60° in ψ. The twofold axes do not coincide with the a, b and diagonal axes but are misaligned by about −4.5° relative to these directions (Fig. 5*a*). In the case of the twinned crystal, the maxima are found at the special values $\psi = 0$, 60, 120° *etc* (Fig. 5*b*). This may be explained if the twin operation is a rotation about the $(a + b)$ diagonal so that the resulting ψ values are the mean of the two components of the twin. (For instance, the 0° value would result from the superposition of the twofold axes located at −4.5 and 4.5°.) There is also a small peak at $\psi = 90^\circ$ and $\varphi = 90^\circ$ (along the c axis) and $\kappa = 180^\circ$, which may result from the similarities between the α- and β-subunits and/or the pseudo-sixfold axis (Chandra *et al.*, 1999).

Table 2
Chiral torsion angles of chromophores found in R-PE of *G. chilensis*.

Angle definitions: σ_{A1}, NC−C4C−CHD−C1D; σ_{A2}, C4C−CHD−C1D−ND *etc.*

Chromophore	φ_{A1}	φ_{A2}	φ_{B1}	φ_{B2}	φ_{C1}	φ_{C2}
1α82	15.65	174.80	−10.14	−6.90	114.08	17.07
1α139	−15.59	−163.71	−20.89	25.24	114.65	−2.34
1β50/61	−26.66	−101.23	1.52	19.14	95.61	28.95
1β82	−13.53	134.30	18.20	9.71	106.36	39.94
1β158	−28.84	154.50	13.74	15.33	117.34	36.42
6α82	21.89	176.96	−5.52	18.92	108.80	9.41
6α139	7.43	−159.16	−15.17	22.94	115.85	12.32
6β50/61	−22.13	−99.37	3.81	9.32	93.73	37.96
6β82	−9.27	138.79	−17.86	19.52	105.72	37.99
6β158	−25.40	164.17	−26.59	7.28	113.87	33.51

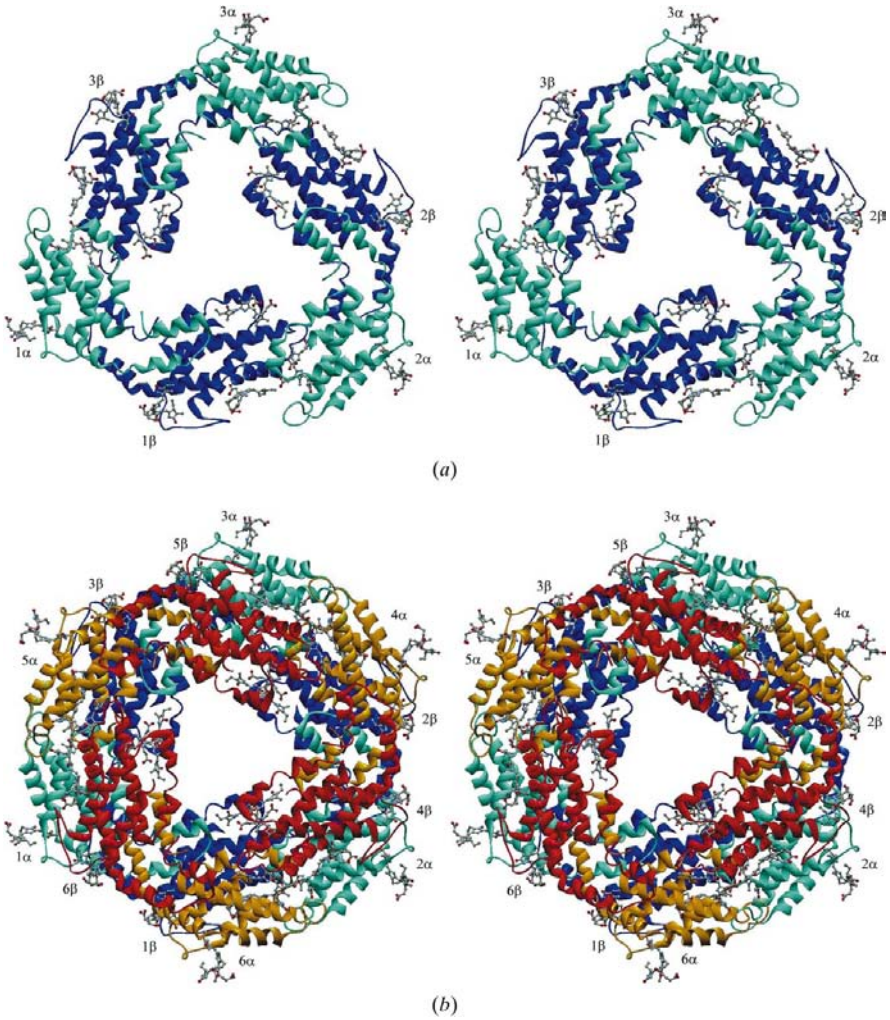

Figure 7
Ribbon stereo representation of (*a*) the trimer ($\alpha_3\beta_3$) (1–3$\alpha\beta$) and (*b*) the hexamer $(\alpha_3\beta_3)_2$ (1–6$\alpha\beta$) as indicated in the figures. Figure prepared with the program *WebLab*3.5 MSI.

3.3. Structure resolution

In the case of a perfect hemihedral twin there is no problem in finding the two possible molecular-replacement solutions and either one may be used (Redinbo & Yeates, 1993; Breyer *et al.*, 1999). For the 2.7 Å resolution untwinned data, *AMoRe* proposed two solutions to the cross-rotation function [Eulerian angles (64, 91, 82°) and (64, 89, 262°)] related by 180° corresponding to the twofold axis relating the two $\alpha\beta$ dimers in the asymmetric unit. In the case of the 2.2 Å resolution twinned data, four solutions were found [Eulerian angles (57, 89, 263°), (57, 91, 83°), (63, 90, 83°) and (63, 91, 263°)], two for each twin domain. The solution with the highest score correlation coefficient of 0.494 and R factor of 0.355 was used for the molecular replacement. Analysis of *SHELX*97 results after the first round of refinement showed that the R-PE crystal is a perfect hemihedral specimen (Table 1).

Because at later stages in refinement only small differences were detected between the side-chain conformations of homologous polypeptide chains, the NCSY approach was used. NCSY significantly reduced the R_{work} and R_{free} factors.

In this work, we have assumed that the primary structure of *G. chilensis* R-PE is very similar to that of R-PE from *Polysiphonia boldii* (Roell & Morse, 1993) (Fig. 1). All the PBPs present a high degree of amino-acid sequence and structure similarities (Fig. 1), with a high α-helix content and the same globin-like folding pattern (Ducret *et al.*, 1994; Apt *et al.*, 1995; Brejc *et al.*, 1995). An extensive sequence-alignment and evolutionary analysis has been performed by Apt *et al.* (1995). Visual inspection of the *G. chilensis* R-PE $(\alpha\beta)_2$ heterodimer cast doubts in a few cases concerning the amino-acid sequence assigned by homology. To confirm the observed amino-acid sequence differences, omit maps were calculated and amino-acid sequence analysis was initiated (Fig. 1). This analysis showed high identity with other reported PEs but also confirmed the changes detected in the electron-density maps.

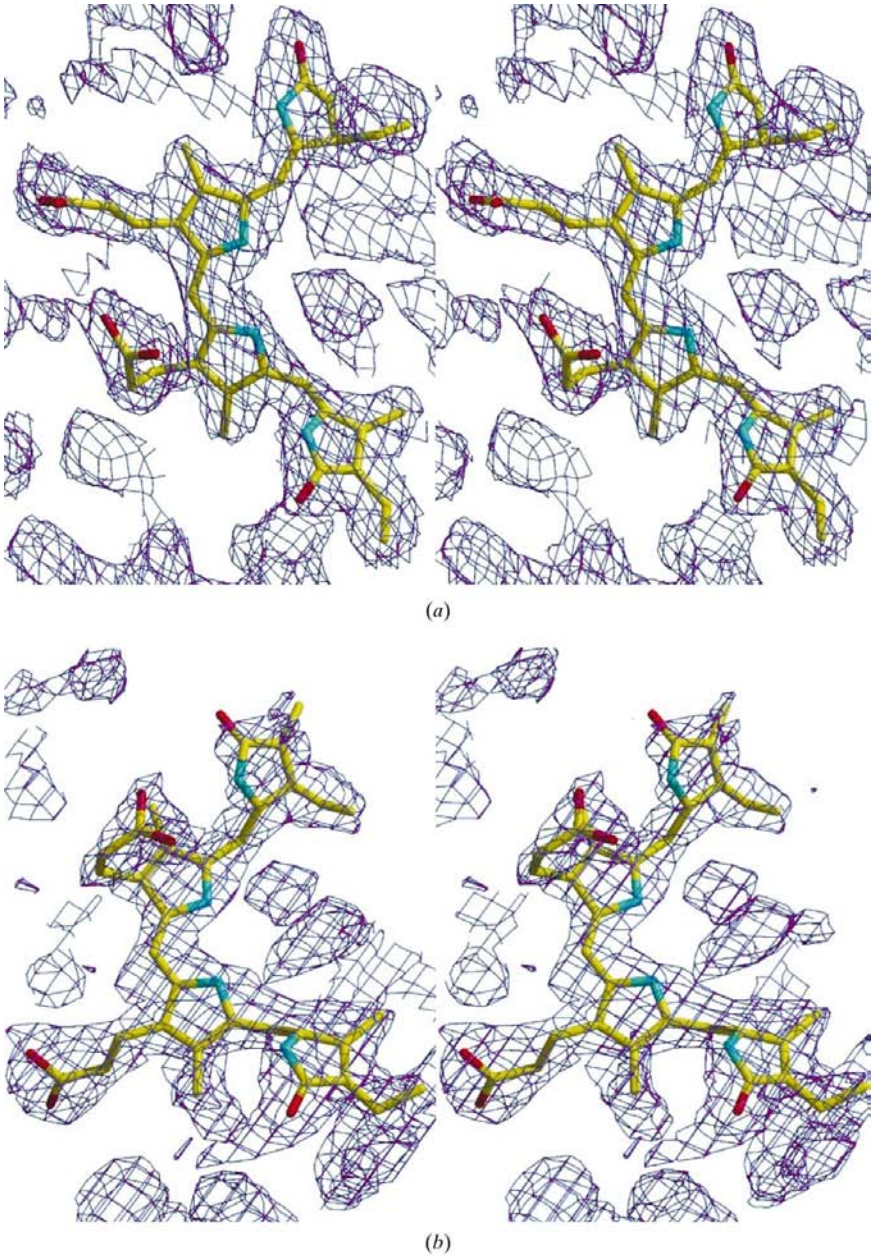

Figure 8
Stereoview of (*a*) α84PEB and (*b*) β50/61PUB chromophore electron-density maps $2F_{obs} - F_{calc}$ contoured at the 1.0σ level. Figure prepared with the programs *XtalView*/*Xfit* v4.0.7 (McRee, 1999*b*) and *Raster3D* v2.5c (Merritt & Bacon, 1997).

3.4. Structure

The α and β polypeptide chains are composed of nine helical segments called *X*, *Y*, *A*, *B*, *E*, *F*, *F*′, *G* and *H* (Fig. 6; Schirmer *et al.*, 1985). In the crystal, there is a hexamer consisting of two $(\alpha_3\beta_3)$ trimers related by a crystallographic twofold axis (Fig. 7). Structure refinement carried out using *SHELX*97 converged to an R_{work} of 0.16 and an R_{free} of 0.25 (Table 1). These values are significantly lower than the R_{work} of 0.34 and R_{free} of 0.35 we obtained with *CNS* 0.3c (Brunger *et al.*, 1998) when the twinning was not taken into account. For statistical reasons, the difference between R_{work} and R_{free} factors can be higher than usual in the presence of a twin (Redinbo & Yeates, 1993; Luecke *et al.*, 1998). Consequently, special care was taken to keep this difference relatively small (Kleywegt & Jones, 1997). Inspection of the electron-density maps indicated that R-PE from *G. chilensis* and PE from *P. boldii* (Roell & Morse, 1993) have the same number of amino acids (164 and 177 in the α- and β-chains, respectively). The stereochemical parameters of the R-PE model fall within the expected values for a well refined structure (residue distribution: most favoured, 93.6%; additional allowed, 6.1%; disallowed, 0.3%; Laskowski *et al.*, 1993; see Table 1). As in other PBSs, only βThr75 falls outside of allowed regions in the Ramachandran plot (not shown); this is because this residue interacts with one of the chromophores. The residues 146–152 in β-chains present very poor electron-density map definition. This area corresponds to an exposed loop in β-chains between the *G* and *H* helixes. The assigned Gly151 is in agreement with the idea of a more flexible loop (highest *B* value) than those reported in 1lia (Val151) and 1b8d (Val151). The commonly reported βAsn72 N-methylation in R-PEs became clear as refinement progressed and Asn72 was modified according to the electron density. As expected, none of the changes detected in the amino-acid sequences affect the

overall folding. The root-mean-square deviations for all main-chain atoms between the R-PEs of *G. chilensis* and *Polysiphonia urceolata* and *Griffithsia monilis* were 2.91 and 2.84 Å, respectively.

The chromophores were easily recognizable in the electron-density maps calculated after initial refinement cycles (Fig. 8). They correspond to two PEB in the α-chain and two PEB plus one PUB in the β-chain (Fig. 6), confirmed by omit maps. In PEB (Fig. 9) the *A*, *B* and *C* rings are co-planar but the *D* rings deviate from the plane. In PUB only the *B* and *C* rings are in the same plane (Table 2). Their configuration, orientation, inter-chromophore distances and the interaction of chromophores with neighboring residues are similar to that found in the R-PEs of *P. urceolata* (Chang *et al.*, 1996) and *G. monilis* (Ritter *et al.*, 1999).

The γ-chain (L_R) is localized in the inner cavity and in the case of R-PE sits on the crystallographic threefold axis. This implies that the γ-chain has only a 1/3 occupancy, which means a weaker density map in the area. However, in our maps some residual electron densities could be recognized close to ring *D* of β82PEB far from the threefold axis. These residual densities indicated the presence of a short α-helix segment and a long lateral chain close to ring *D* of β82PEB with partial occupancies of 0.2–0.6. The 2.2 Å resolution APC structure of *Mastigoclaudus laminosus* was solved from orthorhombic crystals, space group $P2_12_12_1$, and presents a L_C chain of 7.8 kDa (APC-$L_C^{7.8}$) (Reuter *et al.*, 1999). A rigid-body alignment of the APC-$L_C^{7.8}$ (PDB code 1b33) and R-PE of *G. chilensis* indicated the presence of residual density in our map that matched the main α-helix present in the former. Owing to this structural alignment, two 180°-related α-helixes with sequences Ala-Ala-Phe-Arg-Ala-Ala close to the 1β subunit and Ala-Phe-Arg-Ala-Ala close to the 6β subunit could be assigned to the density and refined. The Phe and Arg residues were conserved from the original APC-$L_C^{7.8}$ sequence. The position of Phe is close to that proposed to be a Tyr residue in γ-chain R-PE of *G. monilis*. In the case of R-PE from *P. urceolata,* a residual density close to β82PEB was also modeled as a Phe (Jiang *et al.*, 1999). An amino-acid sequence alignment between APC-$L_C^{7.8}$ and the R-PE *Aglaothamnion neglectum* γ-chain (Apt *et al.*, 1993) did not reveal any similarities. However, a pattern of at least three Phe residues flanked by positively charged residues could be observed.

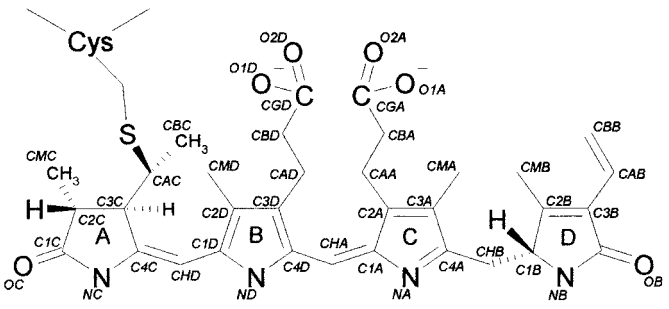

Figure 9
The chemical structure of PEB chromophore. The nomenclature is that used in PDB entry 1cpc and elsewhere (Duerring *et al.*, 1991).

Finally, the environment around ring *D* of β82PEB with Phe and Arg residues is similar to that found close to $\alpha\beta$82PEB except for one missing aromatic residue from the linker.

4. Conclusions

The crystal struture of R-phycoerythrin has been determined in the *R*3 space group. The initial data set at 2.7 Å resolution was used to solve the structure with the molecular-replacement method. Subsequently, a data set to 2.2 Å resolution was collected from a different crystal. Although this crystal also displayed *R*3 symmetry, its diffraction data showed an intensity distribution that indicated merohedral twinning. The higher symmetry of space group *R*32, expected owing to the twinning, was not found because of the slight misalignment of the twin axis relative to the $(a + b)$ diagonal axis. Thus, only a pseudo-sixfold axis was observed in the *hk*3 plane.

The 2.2 Å resolution refined model shows good stereochemistry. The molecule has structural features similar to those described for R-PE of *P. urceolata* (PDB code 1lia) and *G. monilis* (PDB code 1b8d). The interpretation of weak electron densities has led us to propose the presence of a helical segment of linker protein in the central channel of R-PE.

This work has been supported by Dirección de Investigación, Universidad de Concepción grants No. 973172-1 and No. 983177-1, FONDECYT (2970097), the French Committee for Scientific Cooperation with Chile and the National Commission for Scientific and Technological Research (ECOS-CONICYT PC96B02), the CNRS and the CEA. CCM was supported by fellowship from CONICYT and ECOS. We thank Ph. Carpentier, M. Roth and A. Thompson for help at the ID9 and BM14 beamlines of the ESRF, J. Gagnon (LEM/IBS) for protein sequencing and Wladeck Minor (*HKL*2000) and George Scheldrick (*SHELX*97) for assistance.

References

Ajlani, G. & Vernotte, C. (1998). *Eur. J. Biochem.* **25**, 154–159.
Apt, K., Collier, J. & Grossman, A. (1995). *J. Mol. Biol.* **248**, 79–96.
Apt, K., Hoffman, N. & Grossman, A. (1993). *J. Biol. Chem.* **268**, 16208–16215.
Betz, M. (1997). *Biol. Chem.* **378**, 167–176.
Brejc, K., Ficner, R., Huber, R. & Steinbacher, S. (1995). *J. Mol. Biol.* **249**, 424–440.
Breyer, W., Kingston, R., Anderson, B. & Baker, E. (1999). *Acta Cryst.* D**55**, 129–138.
Brunger, A., Adams, P., Clore, M., DeLano, W., Gros, P., Grosse-Kunstleve, R., Jiang, J., Kuszewski, J., Nilges, M., Navraj, P., Read, R., Rice, L., Simonson, T. & Warren, G. (1998). *Acta Cryst.* D**54**, 905–921.
Bunster, M., Tellez, J. & Candia, A. (1997). *Bol. Soc. Chil. Quím.* **42**, 449–455.
Chandra, N., Acharya, K. R. & Moody, P. (1999). *Acta Cryst.* D**55**, 1750–1758.
Chang, W., Jiang, T., Wan, Z., Zhang, J., Yang, Z. & Liang, D. (1996). *J. Mol. Biol.* **262**, 721–731.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ducret, A., Sidler, W., Frank, G. & Zuber, H. (1994). *Eur. J. Biochem.* **221**, 568–580.
Ducret, A., Sidler, W., Werhli, E., Frank, G. & Zuber, H. (1996). *Eur. J. Biochem.* **236**, 1010–1024.
Duerring, A., Schmidt, G. & Huber, R. (1991). *J. Mol. Biol.* **217**, 577–592.
Dumas, P., Ennifar, E. & Walter, P. (1999). *Acta Cryst.* D**55**, 1179–1187.
Esposito, L., Sica, F., Sorrentino, G., Berisio, R., Carotenuto, L., Giordano, A., Raia, C., Rossi, M., Lamzin, V., Wilson, K. & Zagari, A. (1999). *Acta Cryst.* D**54**, 386–390.
Ficner, R. & Huber, R. (1993). *Eur. J. Biochem.* **218**, 103–106.
Ficner, R., Lobeck, K., Schmidt, G. & Huber, R. (1992). *J. Mol. Biol.* **228**, 935–950.
Frazão, C., Sieker, L., Coehlo, R., Morais, J., Pacheco, I., Chen, L., LeGall, J., Dauter, Z., Wilson, K. & Carrondo, M. (1999). *Acta Cryst.* D**55**, 1465–1467.
Gantt, E. (1981). *Ann. Rev. Plant Physiol.* **32**, 327–347.
Gantt, E., Lipschulz, C., Grabowsky, J. & Zimmerman, B. (1979). *Plant Physiol.* **63**, 615–620.
Glazer, A. (1985). *Annu. Rev. Biophys. Biophys. Chem.* **14**, 47–77.
Glazer, A. (1989). *J. Biol. Chem.* **264**, 1–4.
Gomis-Rüth, F., Fita, I., Kiefersauer, R., Huber, R., Aviles, F. & Navaza, J. (1995). *Acta Cryst.* D**51**, 819–823.
Ibuka, A., Taguchi, A., Ishiguro, M., Fushinobu, S., Ishii, Y., Kamitori, S., Okuyama, K., Yamaguchi, K., Konno, M. & Matsuzawa, H. (1999). *J. Mol. Biol.* **285**, 2079–2087.
Ito, N., Komiyama, N. & Fermi, G. (1995). *J. Mol. Biol.* **250**, 648–658.
Jiang, T., Zhang, J.-P. & Liang, D. (1999). *Proteins Struct. Funct. Genet.* **34**, 224–231.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kleywegt, G. & Jones, A. (1997). *Methods Enzymol.* **277**, 208–230.
Koch, E. (1992). *International Tables for Crystallography*, Vol. C, edited by A. J. Wilson, pp. 10–14. Dordrecht: Kluwer.
Kursar, T., van der Meer, J. & Randall, S. (1983). *Plant Physiol.* **73**, 361–347.
Laskowski, R., MacArthur, M., Moss, D. & Thornton, J. (1993). *J. Appl. Cryst.* **26**, 283–291.
Luecke, H., Richter, H. & Lanyi, J. (1998). *Science*, **280**, 1934–1937.
McRee, D. (1999*a*). *Practical Protein Crystallography*, 2nd ed. New York: Academic Press.
McRee, D. (1999*b*). *J. Struct. Biol.* **125**, 156–165.
Matthews, B. (1968). *J. Mol. Biol.* **33**, 491–497.
Merritt, E. & Bacon, D. (1997). *Methods Enzymol.* **277**, 505–524.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Redinbo, M. & Yeates, T. (1993). *Acta Cryst.* D**49**, 375–380.
Reuter, W., Wiegand, G., Huber, R. & Than, M. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 1363–1368.
Ritter, S., Hiller, R., Wrench, P., Welte, W. & Diederichs, K. (1999). *J. Struct. Biol.* **126**, 88–97.
Roell, M. & Morse, D. (1993). *Plant. Mol. Biol.* **21**, 47–58.
Roussel, A. & Cambillau, C. (1989). *Silicon Graphics Geometry Partner Directory*, pp. 77–78. Mountain View, CA, USA: Silicon Graphics.
Schirmer, T., Bode, W., Huber, R., Sidler, W. & Zuber, H. (1985). *J. Mol. Biol.* **184**, 255–277.
Schirmer, T., Huber, R., Schneider, M., Bode, W., Miller, M. & Hackert, M. (1986). *J. Mol. Biol.* **188**, 651–676.
Sheldrick, G. & Schneider, T. (1997). *Methods Enzymol.* **277**, 319–343.
Smith, P., Krohn, R., Hermanson, G., Mallia, A., Gartner, F., Provenzano, M., Fujimoto, E., Goeke, N., Olson, B. & Klenk, D. (1985). *Anal. Biochem.* **150**, 76–85.
Stanley, E. (1972). *J. Appl. Cryst.* **5**, 191–194.
Sweet, R., Fuchs, H., Fisher, R. & Glazer, A. (1977). *J. Biol. Chem.* **252**, 8258- 8260.
Tong, L. & Rossmann, M. (1990). *Acta Cryst.* A**46**, 783–792.
Troxler, R., Ehrhardt, M., Brown-Mason, A. & Offner, G. (1981). *J. Biol. Chem.* **256**, 12176–12184.
Valegård, K., Van Scheltinga, T., Lloyds, M., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Baldwin, J., Schoefield, C., Hajdu. J. & Andersson, I. (1998). *Nature (London)*, **394**, 805–809.
Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.
Wilson, A. (1949). *Acta Cryst.* **2**, 318–321.
Yang, F., Dauter, Z. & Wlodawer, A. (2000). *Acta Cryst.* D**56**, 959–964.
Yeates, T. (1997). *Methods Enzymol.* **276**, 344–358.
Yeates, T. & Fam, B. (1999). *Structure*, **7**, R25-R29.
Yusupov, M., Marquet, W., Ehresmann, C., Ehresmann, B. & Dumas, P. (1999). *Acta Cryst.* D**55**, 281–284.