Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of archaeal glyoxylate reductase from *Pyrococcus horikoshii* OT3 complexed with nicotinamide adenine dinucleotide phosphate

Seiko Yoshikawa,[a]‡ Ryoichi Arai,[a]‡§ Yukiko Kinoshita,[a] Tomomi Uchikubo-Kamo,[a] Taisuke Wakamatsu,[b] Ryogo Akasaka,[a] Ryoji Masui,[b] Takaho Terada,[a] Seiki Kuramitsu,[b] Mikako Shirouzu[a] and Shigeyuki Yokoyama[a,c]*

[a]Protein Research Group, RIKEN Genomic Sciences Center, Tsurumi, Yokohama 230-0045, Japan, [b]Department of Biology, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, and [c]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Bunkyo, Tokyo 113-0033, Japan

‡ These authors contributed equally to this work.
§ Present address: Department of Chemistry, Princeton University, Princeton, NJ 08544-1009, USA.

Correspondence e-mail: yokoyama@biochem.s.u-tokyo.ac.jp

Received 7 August 2006
Accepted 20 December 2006

PDB References: *P. horikoshii* glyoxylate reductase, $I4_1$, 2dbq, r2dbqsf; $P6_1$, 2dbz, r2dbzsf; $P1$, 2dbr, r2dbrsf.

Glyoxylate reductase catalyzes the NAD(P)H-linked reduction of glyoxylate to glycolate. Here, the 1.7 Å crystal structure of glyoxylate reductase from the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3 complexed with nicotinamide adenine dinucleotide phosphate [NADP(H)] determined by the single-wavelength anomalous dispersion (SAD) method is reported. The monomeric structure comprises the two domains typical of NAD(P)-dependent dehydrogenases: the substrate-binding domain (SBD) and the nucleotide-binding domain (NBD). The crystal structure and analytical ultracentrifugation results revealed dimer formation. In the NADP(H)-binding site, the pyrophosphate moiety and the 2′-phosphoadenosine moiety are recognized by the glycine-rich loop (residues 157–162) and by loop residues 180–182, respectively. Furthermore, the present study revealed that *P. horikoshii* glyoxylate reductase contains aromatic clusters and has a larger number of ion pairs and a lower percentage of hydrophobic accessible surface area than its mesophilic homologues, suggesting its thermostability mechanism.

1. Introduction

Glyoxylate reductase (EC 1.1.1.26 or 1.1.1.79) catalyzes the NAD(P)H-linked reduction of glyoxylate to glycolate. It also catalyzes the reduction of hydroxypyruvate to D-glycerate and is thus known to be closely related to hydroxypyruvate reductase (EC 1.1.1.81) and D-glycerate dehydrogenase (EC 1.1.1.29). The first archaeal glyoxylate reductase was found in the hyperthermophilic archaeon *Thermococcus litoralis* DSM 5473 (Ohshima *et al.*, 2001). *T. litoralis* glyoxylate reductase (*Tl*GR) catalyzes the reduction of glyoxylate and hydroxypyruvate; its relative activity for hydroxypyruvate was about a quarter of that for glyoxylate in the presence of NADH as an electron donor. *Tl*GR exhibited low activity for NADPH (7.5% of that for NADH). The enzyme has a fairly high similarity (73% identity) to glyoxylate reductase (PH0597) from the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3 (Kawarabayasi *et al.*, 1998; Fig. 1). *P. horikoshii* glyoxylate reductase (*Ph*GR) consists of 334 amino acids, with a molecular weight of 38 kDa. *Ph*GR has 42% identity to human glyoxylate reductase/hydroxypyruvate reductase (hGRHPR) and 36% identity to D-glycerate dehydrogenase from the methylotrophic bacterium *Hyphomicrobium methylovorum* GM2.

Mutations within the hGRHPR gene are the genetic basis of primary hyperoxaluria type II (PH2), a rare monogenic disorder characterized by a lack of GRHPR activity (Cramer *et al.*, 1999). Several mutations within the GRHPR gene have

been identified in patients with PH2 (Cregeen *et al.*, 2003; Lam *et al.*, 2001; Webster *et al.*, 2000). In hGRHPR, the relative activity for hydroxypyruvate was six times that for glyoxylate in the presence of NADPH (Mdluli *et al.*, 2005). In addition, hGRHPR had higher affinity for NADPH than NADH and exhibited a relatively high activity for NADPH (nearly twice that for NADH). The structure of hGRHPR has recently been reported (Booth *et al.*, 2006).

In contrast, the bacterial D-glycerate dehydrogenase from *H. methylovorum* can reduce hydroxypyruvate and glyoxylate in the presence of NADH and no activity is observed when NADH is substituted by NADPH (Izumi *et al.*, 1990). The crystal structure of the apoenzyme of *H. methylovorum* D-glycerate dehydrogenase (*Hm*GDH) has been reported (Goldberg *et al.*, 1994).

Here, we report the crystal structure of the archaeal glyoxylate reductase from *P. horikoshii* complexed with NADP(H) and discuss aspects of its structure, cofactor recognition and thermostability.

2. Materials and methods

2.1. Protein expression and purification

The glyoxylate reductase (PH0597) gene was amplified by polymerase chain reaction (PCR) from *P. horikoshii* OT3 genomic DNA and was cloned into the plasmid vector pET11a (Novagen). The nonlabelled (native) and selenomethionine-labelled (SeMet) *Ph*GR proteins were expressed in *Escherichia coli* BL21-CodonPlus(DE3)-RIL strain and the methionine-auxotroph *E. coli* B834(DE3)pRARE strain, respectively. The *E. coli* lysate was heated at 343 K for 30 min and the native protein was purified by a series of HiTrap Q, Resource Q, HiTrap Phenyl and Superdex 75 (GE Healthcare) column-chromatography steps. The SeMet-labelled protein was purified by a series of HiTrap Phenyl, HiTrap Q, Resource Q and Superdex 75 column-chromatography steps. The yields of the purified non-labelled *Ph*GR protein and SeMet-labelled protein were 20 and 4 mg per litre of culture, respectively.

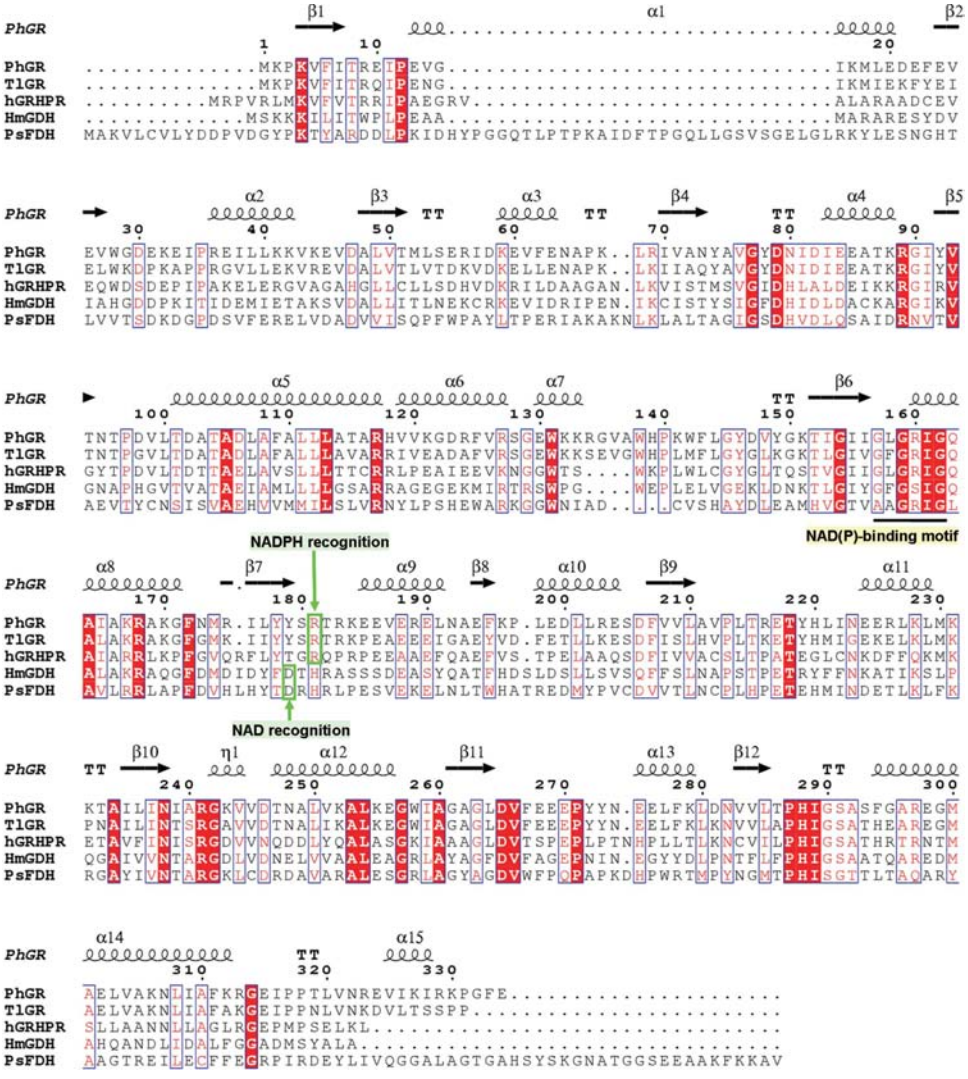

Figure 1
Sequence alignment of glyoxylate reductase homologues. *Ph*GR, *P. horikoshii* glyoxylate reductase; *Tl*GR, *T. litoralis* glyoxylate reductase; hGRHPR, human glyoxylate reductase/hydroxypyruvate reductase; *Hm*GDH, *H. methylovorum* D-glycerate dehydrogenase; *Ps*FDH, *Pseudomonas* sp. formate dehydrogenase. The alignment was generated using *ESPript* (Gouet *et al.*, 1999) and *CLUSTALW* (Thompson *et al.*, 1994). The secondary structures of *Ph*GR ($I4_1$ crystal), as determined by *DSSP* (Kabsch & Sander, 1983), are shown above the sequences.

2.2. Crystallization

Crystals of SeMet *Ph*GR belonging to space group $I4_1$ were obtained in a drop composed of 0.5 µl 11.6 mg ml^{-1} protein solution (20 m*M* Tris–HCl buffer pH 8.0 containing 150 m*M* NaCl, 2 m*M* DTT, 3 m*M* NAD and 5 m*M* α-ketoglutarate) and 0.5 µl reservoir solution (2.0 *M* ammonium sulfate) by the 96-well sitting-drop vapour-diffusion method, equilibrating against 100 µl reservoir solution at 293 K. Diamond-shaped crystals (~200 × 200 × 200 µm) were obtained within a few days.

Crystals of native *Ph*GR belonging to space group $P6_1$ were obtained in a drop composed of 0.5 µl 23.5 mg ml^{-1} protein solution (20 m*M* Tris–HCl buffer pH 8.0 containing 150 m*M* NaCl,

2 m*M* DTT, 3 m*M* NAD and 5 m*M* α-ketoglutarate) and 0.5 µl reservoir solution (0.1 *M* sodium citrate buffer pH 5.6 containing 0.2 *M* potassium sodium tartrate and 2.0 *M* ammonium sulfate) by the 96-well sitting-drop vapour-diffusion method, equilibrating against 100 µl reservoir solution at 293 K. Diamond-shaped crystals ($\sim$300 $\times$ 300 $\times$ 300 µm) were obtained within a few days.

Crystals of native *Ph*GR belonging to space group *P*1 were obtained in a drop composed of 0.5 µl 52.2 mg ml^{-1} protein solution (20 m*M* Tris–HCl buffer pH 8.0 containing 150 m*M* NaCl and 2 m*M* DTT) and 0.5 µl reservoir solution (0.1 *M* Tris buffer pH 8.5 containing 0.2 *M* lithium sulfate and 25% PEG 3350) by the 96-well sitting-drop vapour-diffusion method, equilibrating against 100 µl reservoir solution. Rod-shaped crystals ($\sim$200 $\times$ 100 $\times$ 50 µm) were obtained within a few days at 293 K.

2.3. Data collection, structure determination, refinement and analyses

Data collection was carried out at 100 K, with the reservoir solution containing 25% glycerol (for the $I4_1$ and $P6_1$ crystals) or 15% PEG 400 (for the *P*1 crystal) as a cryoprotectant. The data were collected at BL26B1 (Yamamoto *et al.*, 2002), SPring-8 (Hyogo, Japan) and were recorded on a Jupiter 210 CCD detector (Rigaku, Tokyo, Japan). All diffraction data were processed with the program *HKL*-2000 (Otwinowski & Minor, 1997).

The program *SOLVE* (Terwilliger & Berendzen, 1999) was used for the SeMet crystal (space group $I4_1$) in order to locate the selenium sites and to calculate the phases by the single-wavelength anomalous dispersion (SAD) method and the program *RESOLVE* was used for density modification and partial model building (Terwilliger, 2002). The model was corrected with the program *O* (Jones *et al.*, 1991) and was refined with the programs *LAFIRE* (Yao *et al.*, 2006) and *CNS* (Brünger *et al.*, 1998). The $I4_1$ crystal showed partial hemihedral twinning. The twinning operator was (h, $-k$, $-l$) and the twinning fraction was estimated to be 0.3201 using *CNS* with the input file `detect_twinning.inp`. Structure refinement was accomplished using *CNS* with the input files for refinement with hemihedral twinning.

The structures of the $P6_1$ and *P*1 crystals were solved by the molecular-replacement method using the program *MOLREP* (Vagin & Teplyakov, 1997), with the structure of the $I4_1$ crystal as a search model. The model was corrected iteratively using the program *O* and structure refinement was carried out using the programs *REFMAC*5 (Murshudov *et al.*, 1997), *LAFIRE* and *CNS*.

All refinement statistics are presented in Table 1. The quality of the model was inspected using *PROCHECK* (Laskowski *et al.*, 1993).

The figures were created using the program *PyMOL* (DeLano Scientific, South San Francisco, CA, USA). Superimpositions were carried out using the program *LSQKAB* (Kabsch, 1976) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The ion pairs and the accessible surface area (ASA) were analyzed using the programs *NCONT* and *AREAIMOL* (Lee & Richards, 1971), respectively, from the *CCP*4 suite.

2.4. Analytical ultracentrifugation

The sedimentation-equilibrium experiments were carried out at 293 K using an Optima XL-I analytical ultracentrifuge (Beckman Coulter) with six-channel centrepieces with loading concentrations of 0.68, 0.34 and 0.17 mg ml^{-1}. The sample buffer was 20 m*M* Tris–HCl buffer pH 8.0 containing 150 m*M* sodium chloride and 5 m*M* 2-mercaptoethanol. Data were obtained at 8000, 9000 and 10 000 rev min^{-1}; the absorbance wavelength was 280 nm. The total equilibration time was 16 h for each speed. The estimated partial specific volume of the protein was 0.752 and the estimated solvent density was 1.005 g ml^{-1}. The data were fitted using the manufacturer's software.

2.5. Enzyme assay

Enzymatic activity was determined by monitoring the change in NADH or NADPH absorbance (extinction coefficient = 6.22 mM^{-1} cm^{-1}) at 340 nm using a U-3000 spectrophotometer (Hitachi) at 333 K. All assays were carried out in 100 m*M* potassium phosphate buffer pH 7.5 containing 100 m*M* KCl in a final reaction volume of 1 ml. The concentration of native *Ph*GR was 100 n*M*. The substrate, glyoxylate (20 m*M*) or hydroxypyruvate (3 m*M*), was used with the cofactor NADH or NADPH (0.005–0.2 m*M*). After pre-incubation of the reaction mixture (without enzyme) at 333 K, the reaction was started by addition of the enzyme solution and the decrease in the absorbance at 340 nm was monitored. The data were fitted to a Michaelis–Menten model in order to obtain the Michaelis constant (K_m; for cofactors) and turnover number (k_{cat}).

3. Results

3.1. Structure determination

We determined the structures of three crystals of *Ph*GR belonging to different space groups. The crystals belong to the *I*-centred tetragonal space group $I4_1$, the primitive hexagonal space group $P6_1$ and the primitive triclinic space group *P*1. The $I4_1$ crystal was composed of the SeMet-substituted protein. Its structure was solved by the SAD method and refined to 1.70 Å. There was one *cis*-peptide bond: Glu270–Pro271. The structures of the $P6_1$ and *P*1 crystals (native) were solved by the molecular-replacement method, with the structure of the $I4_1$ crystal as a search model, and were refined to 2.45 and 2.61 Å, respectively. The crystallographic data are summarized in Table 1.

Since *Ph*GR has fairly high similarity (73% identity) to *Tl*GR, we assumed that *Ph*GR would exhibit higher activity for NADH than NADPH in the same way as *Tl*GR (Ohshima *et al.*, 2001). Therefore, 3 m*M* NAD was added to the crystallization solutions in order to crystallize the complex with NAD. The $I4_1$ and $P6_1$ crystals were obtained from crystal-

Table 1
X-ray data-collection, phasing and refinement statistics.

Values in parentheses are for the outer shell.

Crystal	SeMet ($I4_1$)	Native ($P6_1$)	Native ($P1$)
Data collection			
Space group	$I4_1$	$P6_1$	$P1$
Unit-cell parameters (Å, °)	$a = b = 114.25$, $c = 119.60$, $\alpha = \beta = \gamma = 90.00$	$a = b = 104.47$, $c = 182.08$, $\alpha = \beta = 90.00$, $\gamma = 120$	$a = 68.05$, $b = 85.44$, $c = 107.13$, $\alpha = 113.76$, $\beta = 91.17$, $\gamma = 94.00$
Molecules per ASU	1	2	6
Wavelength (Å)	0.97892	1.00000	1.00000
Resolution (Å)	50–1.7 (1.76–1.70)	50–2.45 (2.54–2.45)	50–2.6 (2.69–2.60)
Unique reflections	84158	41156	62135
Redundancy	7.4 (6.5)	7.2 (6.9)	2.6 (2.4)
Completeness (%)	99.9 (99.8)	99.5 (99.8)	94.0 (74.6)
$I/\sigma(I)$	21.9 (4.7)	21.8 (4.5)	14.5 (2.7)
R_{sym} (%)†	8.0 (43.2)	7.2 (46.5)	5.8 (28.8)
SAD analysis			
Resolution (Å)	50.0–1.70		
No. of Se sites‡	5		
FOM_{SAD}§	0.33		
$FOM_{RESOLVE}$¶	0.68		
Refinement			
Resolution (Å)	36.29–1.70	45.31–2.45	49.11–2.61
No. of reflections	83346	37243	62126
No. of protein residues	333	666	1998
No. of water molecules	380	73	82
R_{work} (%)	13.1††	23.9	21.8
R_{free}‡‡ (%)	14.3††	27.6	28.4
R.m.s.d. bond length (Å)	0.005	0.010	0.006
R.m.s.d. bond angles (°)	1.4	1.5	1.2
Average B factor (Å^2)	19.3	62.7	55.0
Ramachandran plot			
Most favoured regions (%)	89.8	87.3	85.5
Additional allowed regions (%)	9.8	12.4	14.0
Generously allowed regions (%)	0.0	0.0	0.2
Disallowed regions (%)	0.3	0.3	0.3
PDB code	2dbq	2dbz	2dbr

† $R_{sym} = \sum |I_i - I_{avg}| / \sum I_i$, where I_i is the observed intensity and I_{avg} is the average intensity. ‡ Number of selenium sites located with *SOLVE*. § Figure of merit after *SOLVE* phasing. ¶ Figure of merit after *RESOLVE*. †† These values were calculated using the twinning operator $(h, -k, -l)$ and the twinning fraction $\alpha = 0.3201$. ‡‡ R_{free} is calculated using a randomly selected 10% of reflections excluded from refinement.

lization conditions containing NAD. However, the electron-density maps obtained from the crystals revealed that the enzyme complexes had formed with NADP(H) instead of NAD. In addition, we tried to crystallize the apoenzyme form under conditions without NAD and obtained the *P*1 crystal. The electron-density map obtained from the *P*1 crystal also revealed enzyme complexed with NADP(H) instead of the apoenzyme. These results indicated that *Ph*GR had formed a stable complex with NADP(H) derived from the expression host *E. coli*. It was unclear whether the cofactor NADP(H) in the crystals was in the reduced or oxidized form.

Since the structure of a homologous enzyme complexed with α-ketoglutarate (a nonphysiological substrate) had been reported (Thompson *et al.*, 2005), we added α-ketoglutarate to the crystallization solutions. However, no electron density corresponding to α-ketoglutarate was present.

3.2. Overall structure

The monomeric structure of *Ph*GR comprises the two domains typical of NAD(P)-dependent dehydrogenases: the substrate-binding domain (SBD; residues 1–99, 292–334) and the nucleotide-binding domain (NBD; residues 100–291) (Fig. 2). The SBD consists of a five-stranded parallel β-sheet (β1–β5) flanked by six α-helices (α1–α4, α14 and α15). The NBD consists of a seven-stranded parallel β-sheet (β6–β12) flanked by nine α-helices (α5–α13). In all of the crystal forms, NADP(H) was bound to the NBD of each monomer. Superimposition of the monomers of the three crystals revealed their structures to be almost the same, except for slight differences in the loop region (29–33) and α9 in the $I4_1$ crystal. The root-mean-square deviations (r.m.s.d.s) are 0.527–0.539 Å ($I4_1$ and $P6_1$ crystals) and 0.455–0.551 Å ($I4_1$ and *P*1 crystals).

In the $I4_1$ crystal, the two *A* chains dimerize by the crystallographic twofold symmetry axis. The *A*–*B* chains in the $P6_1$ crystal and the *A*–*B*, *C*–*D* and *E*–*F* chains in the *P*1 crystal form dimers similar to the dimerization in the $I4_1$ crystal (Fig. 3*a*). According to the results of the analytical ultracentrifugation, the molecular weight of *Ph*GR was 76.5 kDa, indicating that it forms a dimer in solution (the molecular weight of the monomer is 38.0 kDa; Fig. 3*b*). The analytical ultracentrifugation result is consistent with dimer formation in the crystal.

3.3. Comparison with homologous structures

Fig. 4 shows a comparison of the *Ph*GR structure ($I4_1$ crystal) with the structures of its bacterial homologue *Hm*GDH (apoenzyme; PDB code 1gdh; Goldberg *et al.*, 1994) and its human homologue hGRHPR (PDB code 2gcg; Booth *et al.*, 2006). Overall, the structures of *Ph*GR and *Hm*GDH overlap roughly (r.m.s.d. = 2.60 Å). In each domain, the structures of the NBD overlap well (r.m.s.d. = 1.33 Å; Fig. 4*a*) and the structures of the SBD also overlap well (r.m.s.d. = 1.58 Å; Fig. 4*b*). Although the structures of each domain are remarkably conserved, the conformations between the two domains are slightly different. The difference is probably caused by the presence or absence of cofactor. A similar conformation change was found in the structures of *Pseudomonas* sp. formate dehydrogenase (*Ps*FDH) with and without NAD (Lamzin *et al.*, 1994).

Table 2
Enzymatic activity of *Ph*GR.

Substrate	Cofactor	K_m (cofactor) (m*M*)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
Glyoxylate	NADH	0.015	11	733
Glyoxylate	NADPH	0.021	2.4	114
Hydroxypyruvate	NADH	0.013	4.1	315
Hydroxypyruvate	NADPH	0.025	1.8	72

In contrast, in the hGRHPR structure chain *A* is a ternary complex with NADPH and D-glycerate and chain *B* is a binary complex with NADPH. Overall, the structures of *Ph*GR and hGRHPR overlap well (r.m.s.d. = 1.30 Å for chain *A* and 1.38 Å for chain *B*; Fig. 4*c*). In each domain, the structures of both the NBD (r.m.s.d. = 1.21 Å for chain *A* and 1.23 Å for chain *B*) and the SBD (r.m.s.d. = 1.30 Å for chain *A* and 1.36 Å for chain *B*) also overlap well. The conformations between the two domains of *Ph*GR and hGRHPR are almost the same.

3.4. The NADP(H)-binding site

The cofactor, NADP(H), is located within the NBD in the central part of the protein (Fig. 2). In the binding of the nicotinamide moiety, the amide group N7 of the nicotinamide ring is hydrogen bonded to the carboxyl group of Asp265 and the carbonyl group of Ile239 (Fig. 5*a*). Hydrophobic interactions occur between the nicotinamide and the adjacent ribose moieties and several residues such as Val76 and Pro213. The pyrophosphate moiety is recognized by the glycine-rich loop (residues 157–162), which corresponds to the consensus sequence (Gly-*X*-Gly-*X*-*X*-Gly) involved in dinucleotide binding (Wierenga *et al.*, 1985), and by hydrogen bonds to the backbone N atoms of Arg160 and Ile161 and the side-chain N atoms of Arg160. The 2′-phosphoadenosine moiety is recognized by residues 180–182 at the loop region between β7 and α9. The 2′-phosphate group is stabilized by an electrostatic interaction with Arg181 and by a hydrogen bond to Thr182. In addition, a hydrogen bond exists between the 3′-hydroxyl group and Ser180 and hydrophobic interactions occur with the aliphatic moieties of Tyr179, Ser180, Arg181, Val212, Pro213 and Thr218.

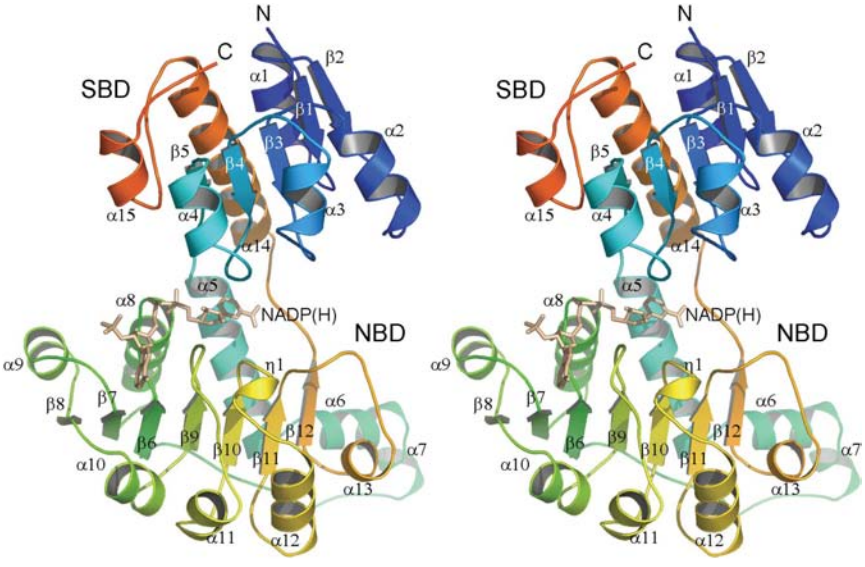

Figure 2
Ribbon representation of the *Ph*GR monomer ($I4_1$ crystal; stereoview). A colour gradient (blue, green, yellow, red) runs from the N-terminus to the C-terminus. NADP(H) is shown as a stick model. SBD, substrate-binding domain; NBD, nucleotide-binding domain.

3.5. Enzymatic activity

Table 2 shows the enzymatic activity of *Ph*GR. The enzyme catalyzed the reduction of both glyoxylate and hydroxypyruvate with either NADH or NADPH as a cofactor. *Ph*GR had lower K_m and higher k_{cat} values for NADH than NADPH with both substrates. The values of the specificity constant k_{cat}/K_m for NADH were 6.4 times (with glyoxylate) and 4.4 times (with hydroxypyruvate) higher than those for NADPH. In addition, the k_{cat} values for glyoxylate were higher than those for hydroxypyruvate with NADH and NADPH.

4. Discussion

4.1. Cofactor specificity

The present study describes the first structure of an archaeal glyoxylate reductase complexed with NADP(H). The 2′-phosphate group of NADP(H) was recognized by Arg181. In the archaeal glyoxylate reductases (*Ph*GR and *Tl*GR) and hGRHPR, which can use either NADPH or NADH, the arginine residue is conserved (Fig. 1). In contrast, in the bacterial *Hm*GDH, which can only use NADH, the arginine residue is replaced by histidine (Fig. 1). In the complex of *Pseudomonas* sp. formate dehydrogenase (*Ps*FDH) with NAD (PDB code 2nad; Lamzin *et al.*, 1994), the 2′- and 3′-hydroxyl groups of the adenosine ribose are recognized by two hydrogen bonds to Asp221. The aspartate residue, which is conserved in *Hm*GDH (Asp177), can bind to the adenosine ribose of NAD, but would prevent NADPH from binding owing to electrostatic repulsion between the negative charges of the aspartate residue and the 2′-phosphate group of NADPH. Since the aspartate residue is substituted by tyrosine and threonine in the archaeal glyoxylate reductases and hGRHPR, respectively, they can utilize either NADPH or NADH as a cofactor. These cofactor-specificity determinants, such as *Ph*GR Arg181 and *Hm*GDH Asp177, are also consistent with those of the shikimate dehydrogenase family (Benach *et al.*, 2003; Michel *et al.*, 2003).

*Ph*GR has a lower K_m value for NADH than for NADPH and exhibits higher activity for NADH than for NADPH (6.4-fold higher with glyoxylate; Table 2), similar to the activity of *Tl*GR (13.3 times higher for NADH than for NADPH with glyoxylate;

Ohshima *et al.*, 2001). The difference in the relative activities of *Ph*GR and *Tl*GR may be caused by the slight variations in their amino-acid sequences, although almost all of the cofactor-binding residues are conserved.

In the crystal structure of *Ph*GR, Arg181, which recognizes the 2′-phosphate group of NADP(H), is located over the adenine moiety of NADP(H), suggesting that the position of Arg181, when interacting with the 2′-phosphate group, interferes with the dissociation and association of NADP(H). When NADP(H) associates with and dissociates from *Ph*GR and the Michaelis–Menten complex with a substrate and NADP(H) is formed during the reaction, some conformational changes around Arg181 may occur to open the NADP(H)-binding site. These conformational changes might confer a relative advantage on NADH over NADPH during the reaction. However, when the reaction is stopped without substrates, *Ph*GR may have a higher affinity for NADPH than for NADH, since Arg181 can recognize the 2′-phosphate group, as suggested by the crystal structure.

hGRHPR has much a lower K_m value for NADPH (K_m = 0.011 m*M*) than for NADH (K_m = 2.42 m*M*) and it exhibits relatively high activity for NADPH (nearly twice that for NADH; Mdluli *et al.*, 2005), while *Ph*GR exhibits higher

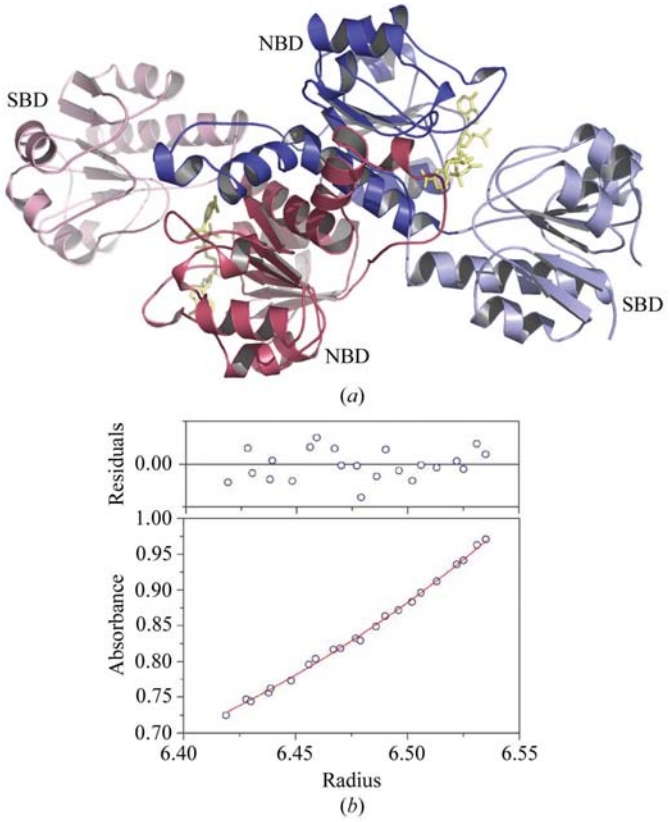

Figure 3
(*a*) Ribbon representation of the dimeric structure of *Ph*GR ($P6_1$ crystal; chain *A*, red; chain *B*, blue). NADP(H) is shown as a stick model. (*b*) A plot of the sedimentation-equilibrium data obtained by analytical ultracentrifugation with the residuals from the best fit to a single ideal species. The bottom panel shows the experimental data (open circles) with the calculated fit (red line). The top panel represents the residuals of the fit. This plot shows data obtained with 0.34 mg ml^{-1} protein and a speed of 9000 rev min^{-1}. The plots represent a single dimeric species model, to which the curve was fitted (MW = 76.5 kDa).

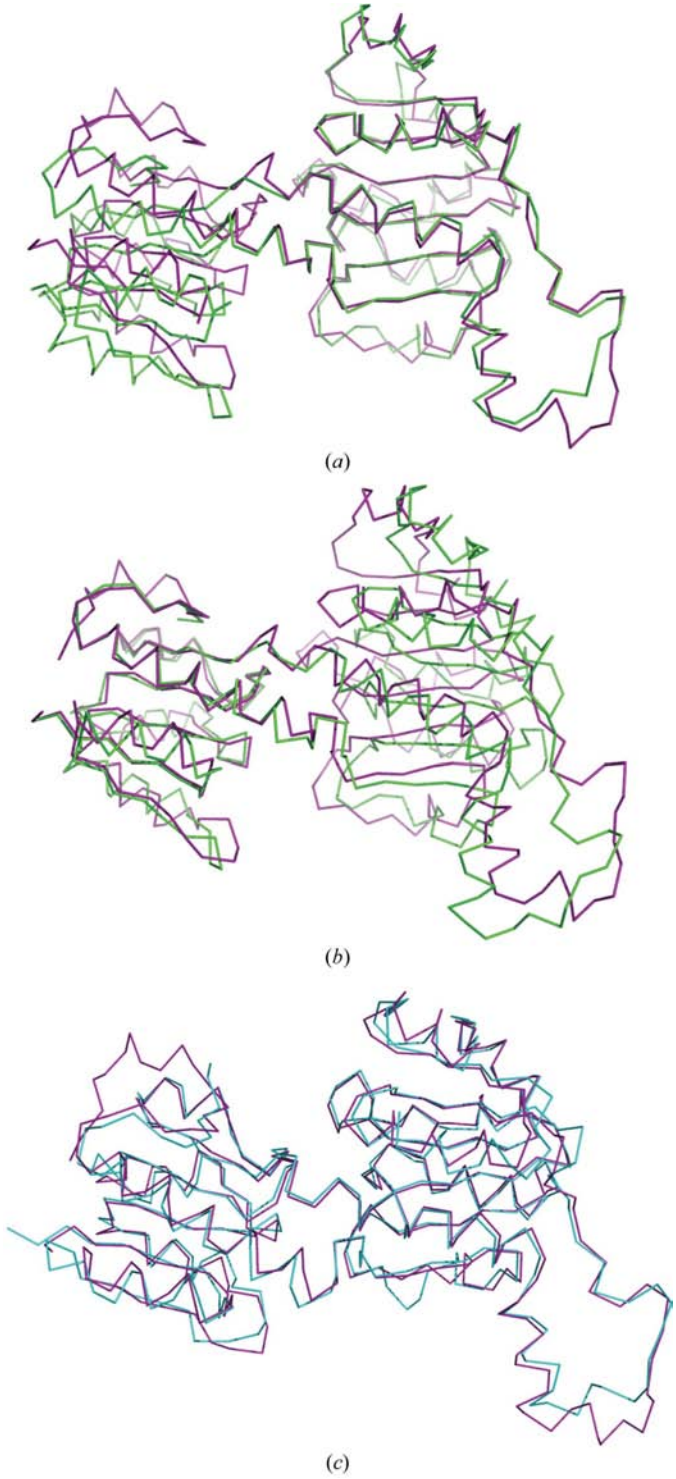

Figure 4
Superimpositions of the main-chain structures of *Ph*GR ($I4_1$ crystal) and its homologues. (*a*) Superimposition of the nucleotide-binding domains (NBD) of *Ph*GR (purple) and *Hm*GDH (green; PDB code 1gdh, chain *A*; r.m.s.d. = 1.33 Å). (*b*) Superimposition of the substrate-binding domains (SBD) of *Ph*GR and *Hm*GDH (r.m.s.d. = 1.58 Å). (*c*) Superimposition of the main-chain structures of *Ph*GR (purple) and hGRHPR (cyan; PDB code 2gcg, chain *A*; r.m.s.d. = 1.30 Å).

activity for NADH than for NADPH. A possible reason for this difference in the cofactor specificity is that the sequences (Fig. 1) and the structures (Fig. 5*b*) of the loop residues 179–184 in *Ph*GR differ from those in hGRHPR (residues 183–188). The recently reported hGRHPR structure revealed that the 2′-phosphate group of NADPH is recognized by two arginine residues, 185 and 188, which may explain the higher affinity for NADPH than for NADH (Booth *et al.*, 2006), while in *Ph*GR the 2′-phosphate group is recognized by one arginine (residue 181).

4.2. Active-site residues

The crystallographic study of *Hm*GDH suggested that the catalytic pair Glu269–His287 and the substrate-orienting Arg240 residue at the active site are involved in the enzymatic reaction (Goldberg *et al.*, 1994). Fig. 5(*c*) shows the superimposition of the active-site residues of *Ph*GR and *Hm*GDH. The Glu–His catalytic pair overlaps well (Glu270–His288 in *Ph*GR). The side-chain orientations of the arginine residues differ between Arg241 in *Ph*GR and Arg240 in *Hm*GDH. In *Ph*GR, a sulfate ion binds to Arg241, implying the potential binding site of a substrate. The distance between the ε-oxygen of Glu270 and the δ-nitrogen of His288 in *Ph*GR ($I4_1$ crystal) is 2.9 Å, suggesting that the Glu270 residue may contribute to the catalytic efficiency by raising the pK_a of His288.

The hGRHPR structure also supports the identification of the active-site residues and the potential substrate-binding site in *Ph*GR. Fig. 5(*d*) shows the superimposition of the active-site residues of *Ph*GR and hGRHPR. The Glu–His catalytic pairs (Glu270–His288 in *Ph*GR; Glu274–His293 in hGRHPR) overlap well in the cases of both the ternary complex (chain *A*) and the binary complex (chain *B*) of hGRHPR. The side-chain orientations of the substrate-orienting arginine residues (Arg241 in *Ph*GR and Arg245 in hGRHPR) are the same. In addition, the substrate, D-glycerate, in the hGRHPR ternary complex structure is located at almost the same site as the sulfate ion in *Ph*GR, indicating the potential binding site of a substrate in *Ph*GR.

4.3. Thermostability

To discuss the thermostability of *Ph*GR from the hyperthermophilic archaeon, we compared the amino-acid compo-

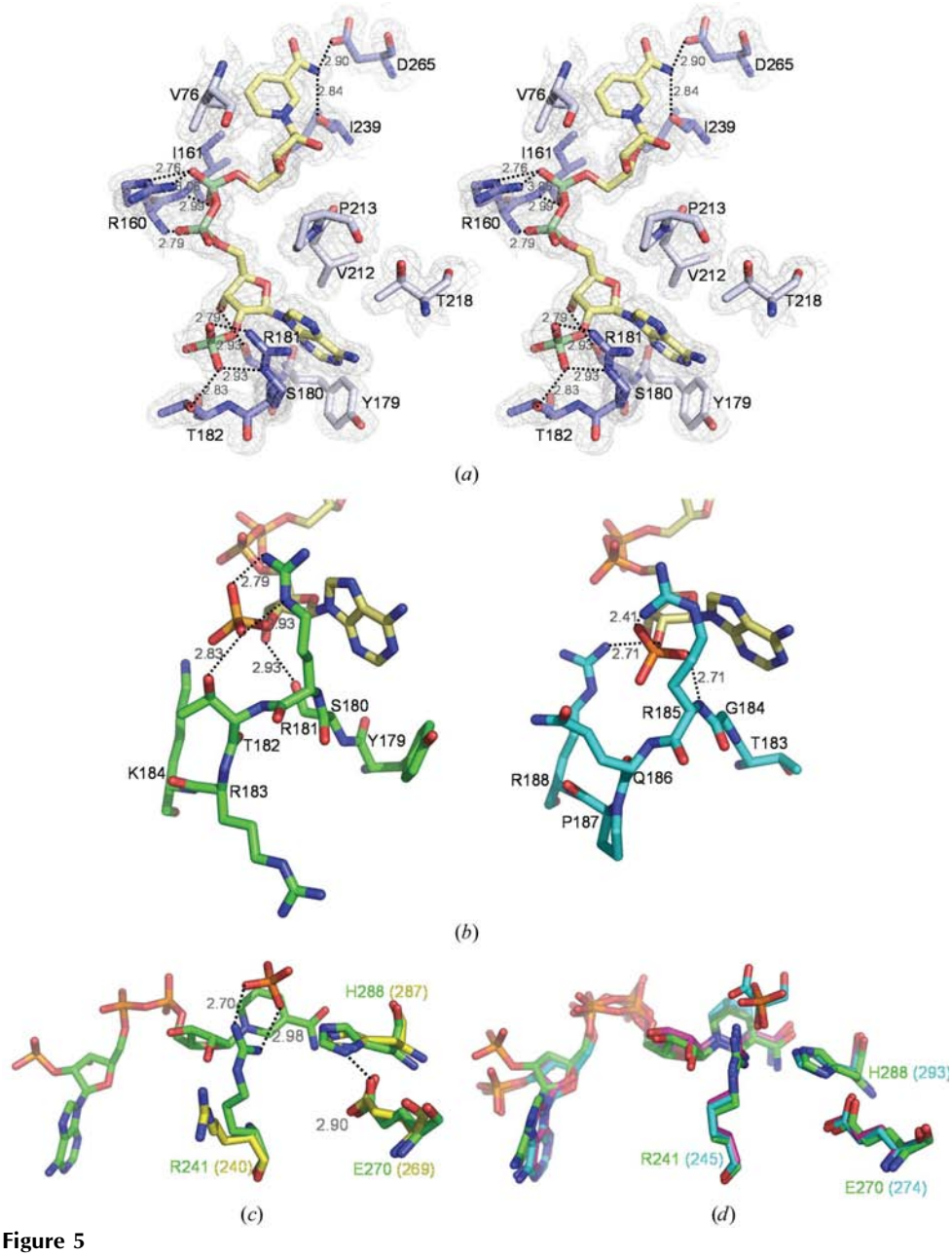

Figure 5
(*a*) Close-up view of stick representation and electron-density map ($2F_o - F_c$) of the NADP(H)-binding site of *Ph*GR ($I4_1$ crystal). Hydrogen bonds and salt bridges are shown as dashed lines together with their distances. The residues involved in hydrogen bonds and salt bridges are coloured blue. The residues involved in hydrophobic interactions are coloured grey. (*b*) Close-up view of the loop regions, which recognize the 2′-phosphate group, at the NADP(H)-binding site in *Ph*GR (left panel; residues 179–184, $I4_1$ crystal) and hGRHPR (right panel, residues 183–188, chain *A*). (*c*) Superimposition of the substrate-binding sites of *Ph*GR ($I4_1$ crystal; green) and *Hm*GDH (chain *A*; yellow). (*d*) Superimposition of the substrate-binding sites of *Ph*GR ($I4_1$ crystal; green) and hGRHPR (chain *A*, cyan, ternary complex with NADPH and D-glycerate; chain *B*, pink, binary complex with NADPH).

Table 3
Characteristics of *Ph*GR, hGRHPR and *Hm*GDH.

	*Ph*GR	hGRHPR	*Hm*GDH
Charged residues† (%)	110 (32.9%)	80 (24.4%)	83 (25.8%)
Neutral residues‡ (%)	69 (20.7%)	94 (28.7%)	94 (29.3%)
Hydrophobic residues§ (%)	155 (46.4%)	154 (46.9%)	144 (44.9%)
Aromatic residues¶ (%)	29 (8.7%)	17 (5.2%)	27 (8.4%)
No. of ion pairs†† (≤4 Å)	67	43	29
Number of ion-pair networks†† (3 residues/4 residues)	12/4	4/2	4/0
Accessible surface area†† (Å^2)			
Overall	27287	25885	24339
Hydrophilic atoms‡‡	14974 (54.9%)	13544 (52.3%)	12549 (51.6%)
Hydrophobic atoms§§	12313 (45.1%)	12341 (47.7%)	11790 (48.4%)

† Charged residues are Asp, Arg, Glu, His and Lys. ‡ Neutral residues are Asn, Cys, Gln, Gly, Ser, Thr and Tyr. § Hydrophobic residues are Ala, Ile, Leu, Met, Phe, Pro, Trp and Val. ¶ Aromatic residues are Phe, Trp and Tyr. †† These values were calculated as dimers (*Ph*GR, $P6_1$ chains *A* and *B*; hGRHPR, 2gcg chains *A* and *B*; *Hm*GDH, 1gdh chains *A* and *B*). ‡‡ Hydrophilic atoms are O, N and S atoms. §§ Hydrophobic atoms are C atoms.

sition, the number of ion pairs and the accessible surface area (ASA) with those of the mesophilic *Hm*GDH and hGRHPR (Table 3). Thermostable proteins frequently show an increased number of charged residues, which lead to increased numbers of ion pairs. Electrostatic interactions play a significant role in stabilizing the protein structure at high temperatures (Yip *et al.*, 1995, 1998; Karshikoff & Ladenstein, 2001; Lebbink *et al.*, 2002). Indeed, *Ph*GR has a greater number of charged residues and ion pairs (≤4 Å) than hGRHPR and *Hm*GDH. The ion-pair networks of *Ph*GR are more expansive than those of hGRHPR and *Hm*GDH and contribute to the structural stabilization of *Ph*GR. There are two significant ion-pair networks consisting of four residues in each monomer of *Ph*GR (Fig. 6*a*). One ion-pair network, consisting of Arg36, Arg56, Asp58 and Glu60, connects α2 and α3, contributing to SBD stabilization. Another ion-pair network, consisting of Arg175, Glu194, Lys196 and Glu204, connects β7, β8 and α10, contributing to NBD stabilization.

*Ph*GR has a larger number of aromatic residues than hGRHPR, while the percentages of hydrophobic residues are almost the same. Closely packed aromatic residues play an important role in protein thermostability (Burley & Petsko, 1985; Vieille & Zeikus, 2001; Li *et al.*, 2003). There are two significant aromatic clusters in *Ph*GR (Fig. 6*b*). The large aromatic cluster, consisting of Phe125, Trp131, Trp142 and Phe143 in one chain (*A* or *B*) and Phe267, Tyr272, Tyr273 and Phe278 in another chain (*B* or *A*, respectively), contributes to the stabilization of the NBD and dimerization. Another aromatic cluster, consisting of Phe23, Tyr92, Phe311 and Phe333, contributes to the stabilization of the SBD. In addition, *Ph*GR has a lower percentage of hydrophobic ASA than hGRHPR and *Hm*GDH. Since surface hydrophobic atoms cannot participate in the stabilization of interactions with solvent molecules, their existence on the protein surface destabilizes protein structures. A number of thermostable proteins have low percentages of hydrophobic ASA (Zhang *et al.*, 2001; Knapp *et al.*, 1999). These structural stabilization factors would play an important role in the thermostability of *Ph*GR.

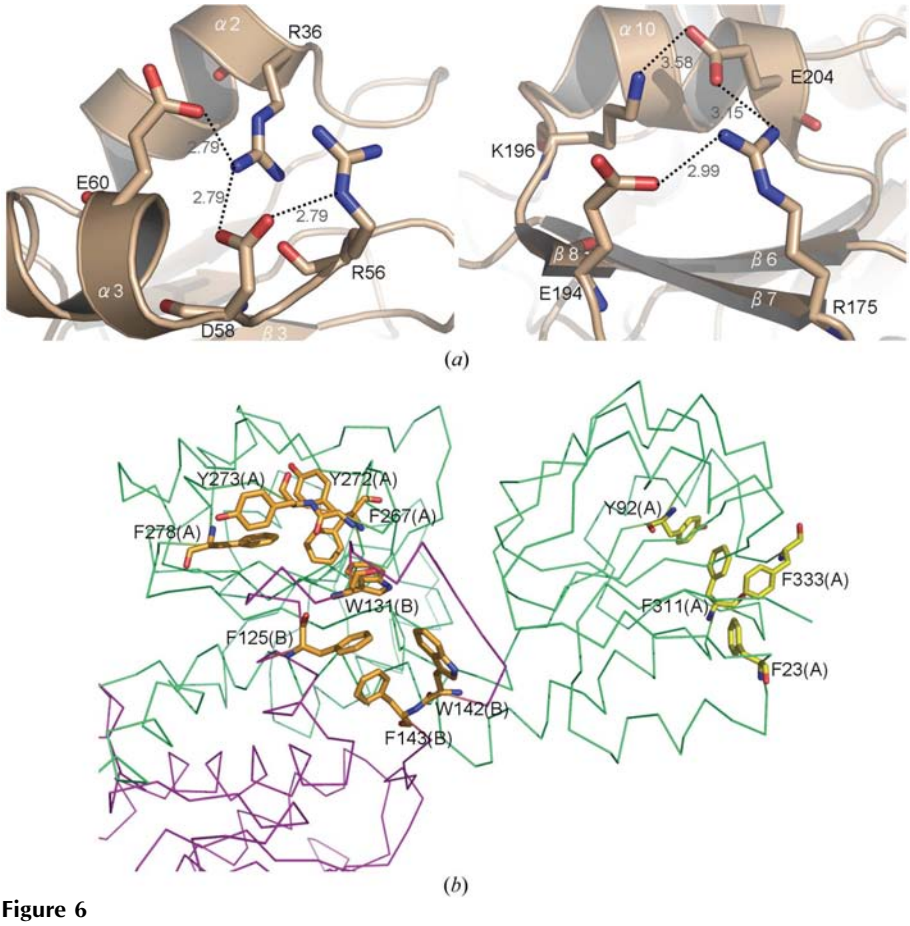

Figure 6
(*a*) Ion-pair networks consisting of four residues in *Ph*GR ($I4_1$ crystal). Left panel, Arg36, Arg56, Asp58 and Glu60; right panel, Arg175, Glu194, Lys196 and Glu204. The residues of the ion-pair networks are shown as stick models on the ribbon-representation structures. (*b*) Aromatic clusters in *Ph*GR ($P6_1$ crystal). The large aromatic cluster (orange) consists of Phe125, Trp131, Trp142 and Phe143 in chain *B* and Phe267, Tyr272, Tyr273 and Phe278 in chain *A*. Another aromatic cluster (yellow) consists of Phe23, Tyr92, Phe311 and Phe333 in chain *A*. The residues of the aromatic clusters are shown as stick models on the main-chain structures (chain *A*, green; chain *B*, purple).

We thank Drs K. Murayama, M. Kukimoto-Niino and T. Matsumoto for helpful advice, Ms Chizu Kuroishi for plasmid preparation, Mr S. Kamo for computer maintenance and Ms A. Ishii, Ms K. Yajima, Ms M. Sunada and Ms T. Nakayama for clerical assistance at RIKEN GSC. We thank Drs K. Kanou, M. Takeda-Shitaka and H. Umeyama of Kitasato University for valuable advice. We also thank Dr M. Yamamoto for

data collection at the RIKEN Structural Genomics beamline BL26B1 at SPring-8. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Benach, J., Lee, I., Edstrom, W., Kuzin, A. P., Chiang, Y., Acton, T. B., Montelione, G. T. & Hunt, J. F. (2003). *J. Biol. Chem.* **278**, 19176–19182.

Booth, M. P., Conners, R., Rumsby, G. & Brady, R. L. (2006). *J. Mol. Biol.* **360**, 178–189.

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.

Burley, S. K. & Petsko, G. A. (1985). *Science*, **229**, 23–28.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cramer, S. D., Ferree, P. M., Lin, K., Milliner, D. S. & Holmes, R. P. (1999). *Hum. Mol. Genet.* **8**, 2063–2069.

Cregeen, D. P., Williams, E. L., Hulton, S. & Rumsby, G. (2003). *Hum. Mutat.* **22**, 497.

Goldberg, J. D., Yoshida, T. & Brick, P. (1994). *J. Mol. Biol.* **236**, 1123–1140.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Izumi, Y., Yoshida, T., Kanzaki, H., Toki, S., Miyazaki, S. S. & Yamada, H. (1990). *Eur. J. Biochem.* **190**, 279–284.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Karshikoff, A. & Ladenstein, R. (2001). *Trends Biochem. Sci.* **26**, 550–556.

Kawarabayasi, Y. *et al.* (1998). *DNA Res.* **5**, 55–76.

Knapp, S., Kardinahl, S., Hellgren, N., Tibbelin, G., Schafer, G. & Ladenstein, R. (1999). *J. Mol. Biol.* **285**, 689–702.

Lam, C. W., Yuen, Y. P., Lai, C. K., Tong, S. F., Lau, L. K., Tong, K. L. & Chan, Y. W. (2001). *Am. J. Kidney Dis.* **38**, 1307–1310.

Lamzin, V. S., Dauter, Z., Popov, V. O., Harutyunyan, E. H. & Wilson, K. S. (1994). *J. Mol. Biol.* **236**, 759–785.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lebbink, J. H., Consalvi, V., Chiaraluce, R., Berndt, K. D. & Ladenstein, R. (2002). *Biochemistry*, **41**, 15524–15535.

Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.

Li, T., Sun, F., Ji, X., Feng, Y. & Rao, Z. (2003). *J. Mol. Biol.* **325**, 1031–1037.

Mdluli, K., Booth, M. P., Brady, R. L. & Rumsby, G. (2005). *Biochim. Biophys. Acta*, **1753**, 209–216.

Michel, G., Roszak, A. W., Sauve, V., MacLean, J., Matte, A., Coggins, J. R., Cygler, M. & Lapthorn, A. J. (2003). *J. Biol. Chem.* **278**, 19463–19472.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Ohshima, T., Nunoura-Kominato, N., Kudome, T. & Sakuraba, H. (2001). *Eur. J. Biochem.* **268**, 4740–4747.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1937–1940.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Thompson, J. R., Bell, J. K., Bratt, J., Grant, G. A. & Banaszak, L. J. (2005). *Biochemistry*, **44**, 5763–5773.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Vieille, C. & Zeikus, G. J. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 1–43.

Webster, K. E., Ferree, P. M., Holmes, R. P. & Cramer, S. D. (2000). *Hum. Genet.* **107**, 176–185.

Wierenga, R. K., De Maeyer, M. C. H. & Hol, W. G. J. (1985). *Biochemistry*, **24**, 1346–1357.

Yamamoto, M., Kumasaka, T., Ueno, G., Ida, K., Kanda, H., Miyano, M. & Ishikawa, T. (2002). *Acta Cryst.* A**58**, C302.

Yao, M., Zhou, Y. & Tanaka, I. (2006). *Acta Cryst.* D**62**, 189–196.

Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D. & Rice, D. W. (1998). *Eur. J. Biochem.* **255**, 336–346.

Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R. & Rice, D. W. (1995). *Structure*, **3**, 1147–1158.

Zhang, X., Meining, W., Fischer, M., Bacher, A. & Ladenstein, R. (2001). *J. Mol. Biol.* **306**, 1099–1114.