Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Improved expression, purification and crystallization of a putative *N*-acetyl-γ-glutamyl-phosphate reductase from rice (*Oryza sativa*)

Jun Miura-Ohnuma,[a]‡ Tsuyoshi Nonaka,[b]‡ Shizue Katoh,[a] Katsuyoshi Murata,[a] Akiko Kita,[c] Kunio Miki[b,d] and Etsuko Katoh[a]*

[a]National Institute of Agrobiological Sciences, Japan, [b]Department of Chemistry, Graduate School of Science, Kyoto University, Japan, [c]Research Reactor Institute, Kyoto University, Japan, and [d]RIKEN Harima Institute/SPring-8, Japan

‡ The first two authors contributed equally to this work.

Correspondence e-mail: ekatoh@nias.affrc.go.jp

Received 7 October 2005
Accepted 30 October 2005
Online 24 November 2005

N-Acetyl-γ-glutamyl-phosphate reductase (AGPR) catalyzes the third step in an eight-step arginine-biosynthetic pathway that starts with glutamate. This enzyme converts *N*-acetyl-γ-glutamyl phosphate to *N*-acetylglutamate-γ-semialdehyde by an NADPH-dependent reductive dephosphorylation. AGPR from *Oryza sativa* (OsAGPR) was expressed in *Escherichia coli* at 291 K as a soluble fusion protein with an upstream thioredoxin-hexahistidine [Trx-$(His)_6$] extension. OsAGPR(Ala50–Pro366) was purified and crystals were obtained using the sitting-drop vapour-diffusion method at 293 K and diffract X-rays to at least 1.8 Å resolution. They belong to the hexagonal space group $P6_1$, with unit-cell parameters a = 86.11, c = 316.3 Å.

1. Introduction

N-Acetyl-γ-glutamyl-phosphate reductase (AGPR; EC 1.2.1.38) catalyzes the third step in an eight-step arginine-biosynthetic pathway that starts with glutamate (Cybis & Davis, 1975). This enzyme converts *N*-acetyl-γ-glutamyl phosphate to *N*-acetylglutamate-γ-semialdehyde (NAGSA) by an NADPH-dependent reductive dephosphorylation. This reaction is reversible and, as NAGSA dehydrogenase, AGPR also catalyzes the reverse reaction, which is the oxidative phosphorylation of NAGSA. Aspartate-β-semialdehyde dehydrogenase (ASADH) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyze the reductive dephosphorylation of β-aspartyl phosphate and the oxidative phosphorylation of glyceraldehyde-3-phosphate using the cofactors NADPH and NAD^+, respectively (Biellmann *et al.*, 1980; Nagradova, 2001). These reactions are mechanistically similar to that of AGPR. The global topologies of ASADH and GAPDH are also similar (Hadfield *et al.*, 1999; Skarzynski *et al.*, 1987). Recently, crystal structures of putative AGPRs from *Thermotoga maritima* (TmAGPR; PDB code 1vkn) and from *Arabidopsis thaliana* (AtAGPR; PDB code 1xyg) have been deposited in the Protein Data Bank. The sequences of AGPRs are similar to those of ASADH and GAPDH; therefore, all three types of enzyme are likely to have similar overall three-dimensional folds. Purification and crystallization procedures for TmAGPR were published in 2003 (Goto *et al.*, 2003). TmAGPR was expressed as the native protein and was purified by a four-step chromatographic procedure. This is the only report concerning the expression, purification and crystallization of an AGPR to date.

Recently, complex plant genomes such as that of *Oryza sativa* (rice; Sasaki & Burr, 2000) have been sequenced. These sequences provide the foundation for plant structural genomic studies, as the human and mouse genome-sequencing projects, as examples, have done for ongoing mammalian structural genomic projects. The rice genome and proteome are suitable and important targets because rice is a major food source and has a relatively small genome in comparison with those of other crops. Furthermore, about 30 000 full-length complementary rice DNA clones exist (Kikuchi *et al.*, 2003) and are publicly available in the Knowledge-Based *Oryza* Molecular Biological Encyclopedia Database (KOME; http://cdna01.dna.affrc.go.jp/cDNA/). Aputative AGPR gene has been found within this database (KOME accession No. AK071544). The corresponding predicted ORF for this AGPR (OsAGPR) is longer than any other known AGPR sequence.

We report optimized conditions for high-level expression, purification and crystallization of the OsAGPR domain (Ala50–Pro366) which is characterized using a combination of protease digestion, matrix-assisted laser desorption-ionization–time-of-flight mass spectrometry (MALDI–TOF MS) and amino-acid sequencing. This is the first report concerning the expression, purification and crystallization of an AGPR from a higher eukaryote.

2. Materials and methods

2.1. Cloning and protein expression

OsAGPR DNA was first synthesized in a PCR reaction mixture using native *Pfu* DNA polymerase (Stratagene), the OsAGPR cDNA (KOME accession No. AK0715443) and the primers the primers 5′-CACC**ATTGAGGGTCGC**ATGGGATCGACGGCGCTCGGT-3′ and 5′-TTAAAGCTTTCAAGGAAATAGGGGTTGATATTGCA-3′, in which the nucleotide sequence in bold corresponds to a Factor Xa cleavage site. Each full-length PCR product was purified by electrophoresis in a 1% agarose gel, isolated using a QIAquick Gel Extraction kit (Qiagen) and inserted into a pENTR/SD/D-TOPO vector using a pENTR/SD/D-TOPO cloning kit (Invitrogen). The sequence fidelity was confirmed by DNA sequencing. Starting with this entry clone, expression vectors were constructed using an LR Reaction kit (Invitrogen). The OsAGPR DNA construct was cloned into an expression vector, pDEST-trx, for expression of the Trx-$(His)_6$/OsAGPR fusion protein (Tsunoda *et al.*, 2005).

For expression of recombinant OsAGPR fusion proteins, *Escherichia coli* BL21 (DE3) competent cells (Novagen) were transformed with each of the expression vectors. Cells were grown at 310 K in 500 ml Luria–Bertani medium containing 50 µg ml^{-1} ampicillin (LB^{amp}) to an absorbance of approximately 0.6 at 600 nm. Expression was induced by addition of isopropyl-1-thio-β-galactopyranoside (IPTG) to a final concentration of 0.1 m*M*. Cells were grown for an additional 16 h at 291 K, centrifuged at 5000*g* for 15 min and then frozen.

2.2. Purification of OsAGPR(Ala50–Pro366)

A frozen cell pellet obtained from a 500 ml LB^{amp} culture was suspended in 100 ml 20 m*M* Tris–HCl pH 7.8, 0.5 *M* NaCl, 2.5 m*M* 2-mercaptoethanol (2-ME), 10 m*M* imidazole, 1.0 m*M* phenylmethanesulfonyl fluoride (PMSF) and disrupted by sonication. Lysed cell suspensions were centrifuged at 70 000*g* for 20 min at 277 K.

Soluble Trx-$(His)_6$/OsAGPR, which was present in the supernatant of the centrifuged sample, was loaded at a flow rate of 2 ml per minute directly onto a 5 ml HiTrap chelating HP column (Amersham Biosciences) bound with Ni (Ni-chelating column) and pre-equilibrated with the buffer used for sonication. The column was exhaustively washed with 20 m*M* Tris–HCl pH 7.8, 0.5 *M* NaCl, 1.0 m*M* 2-ME, 10 m*M* imidazole and the recombinant protein was eluted with a linear gradient of 10 m*M* to 0.5 *M* imidazole. The eluate containing imidazole was chromatographed over HiLoad 26/60 Superdex 200 pg (Amersham Biosciences) in the presence of 20 m*M* Tris–HCl pH 7.8, 0.1 *M* NaCl, 2.5 m*M* 2-ME. The Trx-$(His)_6$ tag and the first 49 residues of OsAGPR were removed by cleavage with 20 U Factor Xa (Novagen) per milligram of protein at 310 K for 16 h. The solution, which contained the truncated form of OsAGPR, referred to hereafter as OsAGPR(Ala50–Pro366), was applied onto a 5 ml Ni-chelating column at a flow rate of 2 ml min^{-1}. The column had been pre-equilibrated with 20 m*M* Tris–HCl pH 7.8, 0.1 *M* NaCl, 1.0 m*M* 2-ME, 10 m*M* imidazole. The column was exhaustively washed with the same buffer and OsAGPR(Ala50–Pro366) eluted using a linear gradient of 10 m*M* to 0.5 *M* imidazole. The appropriate protein fraction was then chromatographed over HiLoad 26/60 Superdex 200 pg in 20 m*M* Tris–HCl pH 7.8, 0.1 *M* NaCl, 2.5 m*M* 2-ME.

2.3. Crystallization and X-ray data collection of OsAGPR(Ala50–Pro366)

The protein solution was dialyzed against 10 m*M* Tris–HCl pH 7.5 and concentrated to approximately 10 mg ml^{-1} by centrifugal ultrafiltration. OsAGPR(Ala50–Pro366) crystals were obtained by the sitting-drop vapour-diffusion method at 293 K. Mixtures of the protein solution and the same volume of reservoir solution, composed of 0.1 *M* Tris–HCl pH 8.5, 0.72 *M* sodium formate and 22.5% polyethylene glycol monomethyl ether 2000, were equilibrated against the reservoir solution.

The crystal was flash-frozen under a cryogenic nitrogen-gas stream. An X-ray diffraction data set was collected from the crystal at the SPring-8 BL44B2 beamline at 90 K at a wavelength of 1.0080 Å. The crystal-to-detector distance was set to 240 mm and diffraction images were recorded on an ADSC Quantum 210 CCD detector with 0.3° oscillation and 3 s exposure per frame. A total of 600 frames were collected with an oscillation range of 180°. The corrected data set was processed using the *HKL*2000 suite (Otwinowski & Minor, 1997).

3. Results and discussion

For OsAGPR expression in *E. coli*, we constructed three Gateway expression vectors. One system, incorporating the pDEST14 vector, was designed to express native OsAGPR, while the other systems, incorporating the pDEST17 or the pDEST-trx vector, were designed to express OsAGPR fused to hexahistidine [$(His)_6$] or OsAGPR fused to thioredoxin-hexahistidine [Trx-$(His)_6$], respectively. We first tried to express native OsAGPR using 1.0 m*M* IPTG and the pDEST14 system, knowing that it had previously been possible to directly express high levels of native TmAGPR (Goto *et al.*, 2003). However, the pDEST14 system did not produce recombinant OsAGPR when the cells were cultured at 291 or 310 K (data not shown). In general, proteins from thermophilic organisms, *e.g. T. maritima*, are much more thermostable and thus easier to purify and characterize (Lesley *et al.*, 2002) than are non-thermostable proteins. Additionally, proteins of higher eukaryotes are typically difficult to express in bacteria. Although the OsAGPR and TmAGPR amino-acid sequences are similar (40.5% identity; see Supplementary Figure S1[1]), this similarity is not reflected in their recombinant expression patterns. We next tried to express $(His)_6$/OsAGPR. Although we were able to express large and approximately equivalent amounts of $(His)_6$/OsAGPR at both concentrations of IPTG (1.0 and 0.1 m*M*) and at both temperatures (310 and 291 K), the protein always precipitated after sonication (Supplementary Figure S2, lanes 3, 4, 7, 8, 11 and 12). More Trx-$(His)_6$/OsAGPR was produced using 1.0 m*M* IPTG at 310 K than at 291 K, but all protein produced at the higher temperature precipitated after sonication (Supplementary Figure S2, lanes 15 and 16). Trx-$(His)_6$/OsAGPR was also induced using 0.1 m*M* IPTG at 310 K and in this case about 10% of the fusion protein was soluble after sonication (Supplementary Figure S2, lanes 19 and 20). Furthermore, at the lower temperature (291 K) and using a relatively small amount of IPTG for induction (0.1 m*M*), approximately 30% of the Trx-$(His)_6$/OsAGPR remained soluble after sonication (Supplementary Figure S2, lanes 23 and 24). These results

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference PU5116).

may be explained by the decreased rate of expression and the decreased accumulation of the exogenous protein *in vivo* that is brought about by the lower temperature and the relatively small amount of IPTG used to induce protein expression, in conjunction with the solubilizing effects of thioredoxin. Indeed, it has been our experience that in most cases recombinant rice proteins fused to an upstream thioredoxin construct are significantly more soluble and are expressed at higher levels than are their unmodified counterparts or if they are fused to other extensions (Tsunoda *et al.*, 2005). LaVallie and coworkers also found that heterologous proteins fused to thioredoxin and targeted to the *E. coli* cytoplasm accumulated to high levels and were usually soluble (LaVallie *et al.*, 1993).

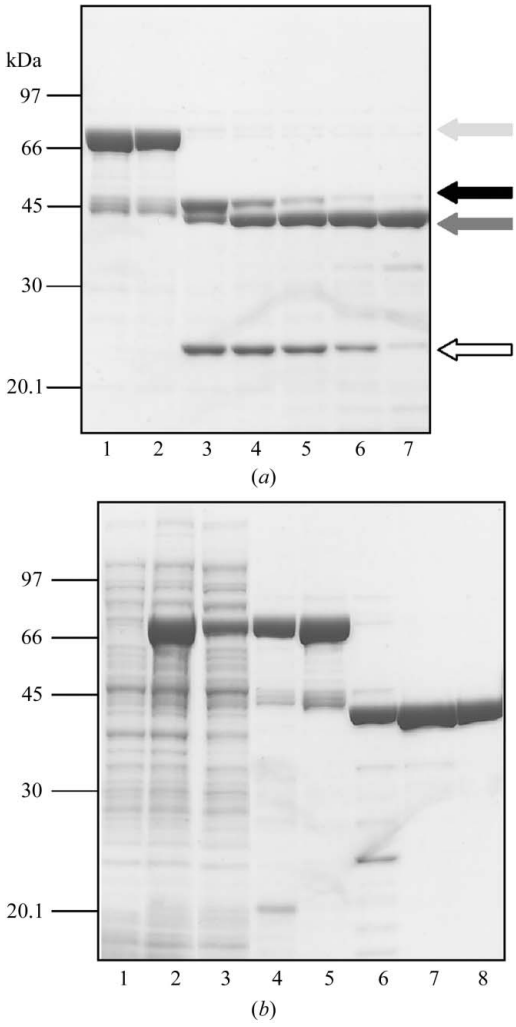

Figure 1
Coomassie Brilliant Blue stained SDS–PAGE gel (12.5% acrylamide). (*a*)Trx-$(His)_6$/OsAGPR before proteolysis by Factor Xa (lane 1) and after 16 h of Factor Xa digestion at concentrations of 0, 0.01, 0.05, 0.1, 0.2 and 0.5 U per 10 mg of Trx-$(His)_6$/OsAGPR (lanes 2–7, respectively). The light grey arrow marks the position of Trx-$(His)_6$/OsAGPR. The black arrow marks the position of the larger form of OsAGPR, which has a molecular weight of approximately 45 kDa. The dark gray arrow marks the position of the smaller form of OsAGPR [OsAGPR(Ala50–Pro366)], which has a molecular weight of approximately 40 kDa. The open arrow marks the position of the Trx-$(His)_6$ tag. (*b*) OsAGPR at various stages of expression and purification. Lane 1, the uninduced BL21(DE3) whole cell extract; lane 2, the BL21 (DE3) cell extract after 16 h of induction at 291 K; lane 3, the soluble fraction of induced BL21 (DE3) cell lysate; lane 4, Trx-$(His)_6$/OsAGPR after Ni-chelating chromatography; lane 5; Trx-$(His)_6$/OsAGPR that eluted as a monomer upon gel filtration; lane 6, OsAGPR(Ala50–Pro366) after Factor Xa digestion; lane 7; OsAGPR(Ala50–Pro366) after Ni-chelating chromatography; lane 8; OsAGPR(Ala50–Pro366) after Superdex gel filtration.

We purified Trx-$(His)_6$/OsAGPR that had been recovered in the soluble fraction. While the fusion protein remained soluble, about 60% eluted in the volume expected for the monomeric protein. Since the void volume corresponds to a molecular weight much greater than that expected for monomeric Trx-$(His)_6$/OsAGPR, the protein had probably aggregated. On the other hand, the protein that eluted as if it was monomeric during gel chromatography could be digested with Factor Xa. An SDS–PAGE gel of the digested protein shows two major bands (black and dark grey arrow, Fig. 1*a*) other than the Trx-$(His)_6$ fragment (open arrow, Fig. 1*a*) whose relative intensities depend on the amount of Factor Xa used in the digestion. When Trx-$(His)_6$/OsAGPR was digested with more than 0.2 U Factor Xa per 10 mg fusion protein, only the smaller (approximately 40 kDa) of the two protein bands was present (dark grey arrow; Fig. 1*a*, lanes 6 and 7). The molecular weight of the upper band (about 45 kDa) is nearly the same as that calculated for the predicted ORF OsAGPR sequence (44.8 kDa).

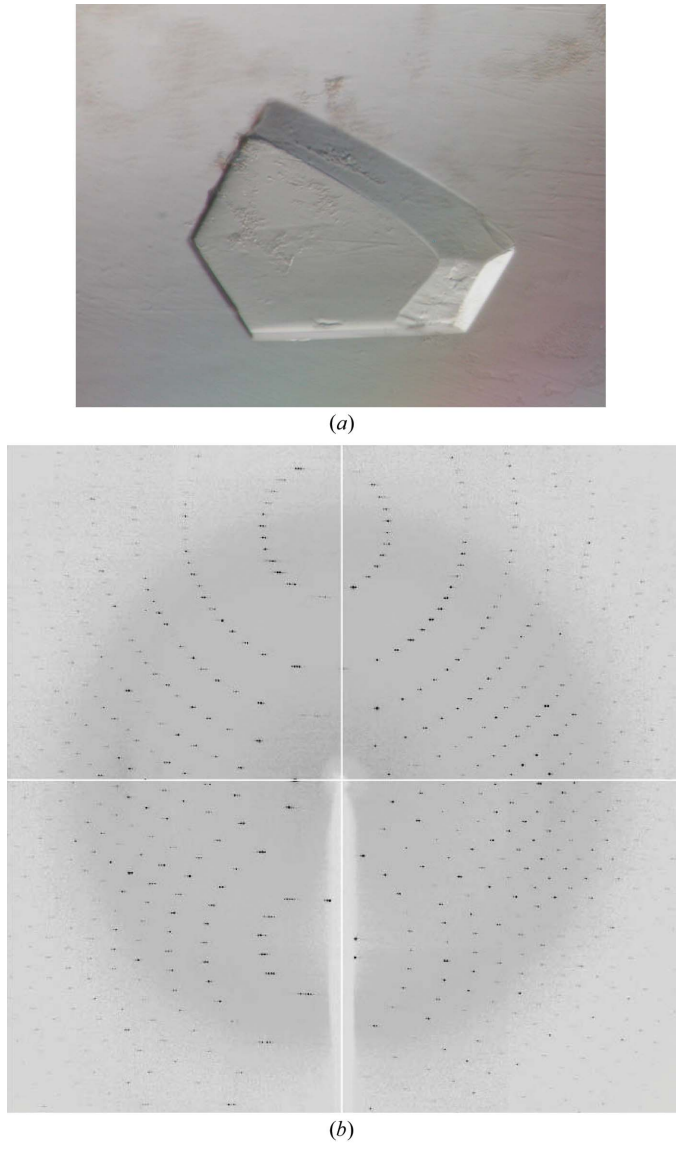

Figure 2
Crystallization of OsAGPR(Ala50–Pro366). (*a*) A typical crystal with dimensions of 0.5 × 0.3 × 0.1 mm. (*b*) X-ray diffraction pattern diffracting beyond 1.8 Å resolution.

Table 1
Summary of the data collection of OsAGPR.

Values in parentheses are for the outermost resolution shell.

Resolution range (Å)	50–2.20 (2.28–2.20)
Space group	$P6_1$
Unit-cell parameters	
a (Å)	86.11
c (Å)	316.3
Wavelength (Å)	1.0080
Temperature (K)	100
$R_{\rm merge}$† (%)	4.2 (6.4)
Completeness (>1σ) (%)	97.2 (88.8)
No. of observed reflections (>1σ)	667362
No. of unique reflections (>1σ)	65608
Overall redundancy (>1σ)	10.2
$I/\sigma(I)$	50.8 (23.7)

† $R_{\rm merge} = \sum |I_i - \langle I_i \rangle| / \sum \langle I_i \rangle$, where I_i is the observed intensity and $\langle I_i \rangle$ is the average intensity over symmetry-equivalent measurements.

We could not control the Factor Xa digestion to the extent that the major product was the larger molecular-weight form of OsAGPR. However, as the smaller of the two proteins resisted further proteolysis (Fig. 1*a*, lanes 6 and 7), we exploited this property to isolate and characterize the protein using MALDI–TOF MS and N-terminal amino-acid sequencing (Supplementary Figure S3). From these analyses, the faster migrating of the two bands, referred to hereafter as OsAGPR(Ala50–Pro366), is a form of OsAGPR that is missing the first 49 predicted ORF residues. At this stage of the investigation, we believed that this foreshortened OsAGPR would be suitable for X-ray crystallography because it appeared to be a stable and folded domain as demonstrated by its resistance to Factor Xa proteolysis (Fig. 1*a*). It may be that the region preceding residue Ala50 is flexible: only a small amount of Factor Xa was required to remove it. It is curious that this N-terminal region is not found for the other two AGPRs of known sequence (Supplementary Figure S1). Perhaps the first 49 residues, while part of a single domain that includes Ala50–Pro366, are linked to the remainder of the protein primarily by a solvent-accessible and flexible loop. Alternatively, it may be that the ORF prediction is wrong and only a shorter form of OsAGPR is produced in rice. Fig. 1(*b*) summarizes the results in each purification step.

The crystals grew in about 10 d to approximate dimensions of 0.5 × 0.3 × 0.1 mm (Fig. 2). Crystals were picked up in a nylon loop and successfully flash-frozen in a cryogenic nitrogen-gas stream. The crystals diffract X-rays beyond 1.8 Å resolution, but diffraction data to less than 2.20 Å resolution were collected and processed owing to the large length of the c axis. They belong to the hexagonal space group $P6_1$, with unit-cell parameters a = 86.11, c = 316.3 Å. Assumption of the presence of four OsAGPR molecules in the asymmetric unit gave a reasonable Matthews coefficient $V_{\rm M}$ of 2.2 Å^3 Da^{-1} and a solvent content of 45% (Matthews, 1968). Statistics of the data collection are summarized in Table 1.

This work was supported in part by grants PR-4101 and IP-6001 from the Rice Genome Project sponsored by the Ministry of Agriculture, Forestry and Fisheries of Japan to EK.

References

Biellmann, J. F., Eid, P., Hirth, C. & Jornvall, H. (1980). *Eur. J. Biochem.* **104**, 53–58.
Cybis, J. & Davis, R. H. (1975). *J. Bacteriol.* **123**, 196–202.
Goto, M., Agari, Y., Omi, R., Miyahara, I. & Hirotsu, K. (2003). *Acta Cryst.* D**59**, 356–358.
Hadfield, A., Kryger, G., Ouyang, J., Petsko, G. A., Ringe, D. & Viola, R. (1999). *J. Mol. Biol.* **289**, 991–1002.
Kikuchi, S. *et al.* (2003). *Science*, **301**, 376–379.
LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. & McCoy, J. M. (1993). *Biotechnology*, **11**, 187–193.
Lesley, S. A. *et al.* (2002). *Proc. Natl Acad. Sci. USA*, **99**, 11664–11669.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nagradova, N. K. (2001). *Biochemistry (Mosc.)*, **66**, 1067–1076.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sasaki, T. & Burr, B. (2000). *Curr. Opin. Plant Biol.* **3**, 138–141.
Skarzynski, T., Moody, P. C. & Wonacott, A. J. (1987). *J. Mol. Biol.* **193**, 171–187.
Tsunoda, Y., Sakai, N., Kikuchi, K., Katoh, S., Akagi, K., Miura-Ohnuma, J., Tashiro, Y., Murata, K., Shibuya, N. & Katoh, E. (2005). *Protein Expr. Purif.* **42**, 268–277.