Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Substrate recognition and selectivity of plant glycerol-3-phosphate acyltransferases (GPATs) from *Cucurbita moscata* and *Spinacea oleracea*

Taro Tamada, Michael D. Feese, Stefano R. Ferri, Yoichi Kato, Rieko Yajima, Toshihiro Toguri and Ryota Kuroki*

Central Laboratories for Key Technology, Kirin Brewery Co. Ltd, Fukuura 1-13-5, Kanazawa-ku, Yokohama 236-0004, Japan

Correspondence e-mail: r-kuroki@kirin.co.jp

Received 16 April 2003
Accepted 19 September 2003

PDB Reference: glycerol-3-phosphate acyltransferase, 1iuq, r1iuqsf.

Stromal glycerol-3-phosphate acyltransferases (GPAT) are responsible for the selective incorporation of saturated and unsaturated fatty-acyl chains into chloroplast membranes, which is an important determinant of a plant's ability to tolerate chilling temperatures. The molecular mechanisms of plant chilling tolerance were elucidated by creating chimeric GPATs between squash (*Cucurbita moscata*, chilling-sensitive) and spinach (*Spinacea oleracea*, chilling-tolerant) and the results were interpreted using structural information on squash GPAT determined by X-ray crystallography at 1.55 Å resolution. Enzymatic analysis of the chimeric GPATs showed that the chimeric GPATs containing the spinach region from residues 128 to 187 prefer the 18:1 unsaturated fatty acid rather than 16:0 saturated fatty acid. Structure analysis suggests that the size and character of the cavity that is formed from this region determines the specific recognition of acyl chains.

1. Introduction

Many plant species growing in torrid and subtorrid zones are injured or killed by exposure to low non-freezing temperatures (273–288 K). These species, which include many economically important crops such as rice, maize, banana, soybean and squash, are classified as chilling-sensitive plants. On the other hand, temperate plants, such as barley, lettuce, field pea and spinach, are tolerant of such chilling temperatures. An important factor in the ability of plants to tolerate chilling temperature has long been thought to be membrane fluidity (Lyons, 1973), which in turn is greatly influenced by the degree of unsaturation of the membrane lipids. More specifically, a correlation between chilling sensitivity and the presence of disaturated plastidial phosphatidyl glycerol (PG) has been found in a large number of plants (Murata *et al.*, 1982; Murata, 1983; Roughan, 1985; Bishop, 1986); that is to say, chilling-sensitive plants have a high proportion of saturated fatty acids [palmitic acid (16:0), stearic acid (18:0) and 3-*trans*-hexadecenoic acid (16:1), which has similar physical properties to 16:0] in the PG of the plant's chloroplast membranes. In contrast, chilling-tolerant plants have a high proportion of *cis*-unsaturated fatty acids [oleic acid (18:1), linoleic acid (18:2) and α-linoleic acid (18:3)] in their PG.

Fatty-acid biosynthesis in plants occurs in the chloroplast, where it proceeds by the sequential addition of two-carbon units to the acyl chain attached to the acyl-carrier protein (ACP; Ohlrogge & Browse, 1995; Harwood, 1996). Stromal glycerol-3-phosphate acyltransferase (GPAT; EC 2.3.1.15) catalyses the first committed step of this biosynthesis (Frentzen, 1993): transfer of the acyl chain from mainly

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

Table 1
Data collection and phase refinement.

Values in parentheses refer to the high-resolution shell.

	SPring-8 (BL40B2)			Photon Factory (BL18B)		
X-ray source	Edge	Peak	Remote	Edge	Peak	Remote
Data collection						
Wavelength (Å)	0.9781	0.9778	0.9840	0.9796	0.9793	0.9700
Resolution (Å)	2.0	2.0	2.0	1.9	1.9	1.9
Unique reflections	29589	29633	29541	34714	34713	34713
R_{sym}†	0.061 (0.278)	0.064 (0.258)	0.037 (0.113)	0.060 (0.232)	0.062 (0.244)	0.071 (0.368)
Completeness (%)	99.8 (99.5)	99.9 (99.9)	99.7 (99.7)	100.0 (100.0)	100.0 (100.0)	99.9 (99.9)
Redundancy	3.8 (3.7)	7.0 (6.8)	3.8 (3.8)	7.2 (7.2)	7.2 (7.2)	7.2 (7.3)
$I/\sigma(I)$	4.2 (0.8)	9.7 (2.8)	14.4 (6.3)	7.3 (2.9)	7.0 (3.0)	6.9 (2.0)
R_{anom}‡	0.059 (0.061)	0.053 (0.127)	0.025 (0.079)	0.036 (0.101)	0.049 (0.118)	0.044 (0.159)
Phase refinement (at 2.5 Å)						
R_{iso}§	–	0.022	0.034	0.036	0.038	0.047
Phasing power¶						
Centric	–	0.46	0.88	0.46	0.64	0.96
Acentric	–	0.68	1.29	0.70	0.99	1.46
R_{cullis}††						
Centric	–	0.91	0.75	0.94	0.86	0.74
Acentric	0.72	0.90	0.77	0.90	0.82	0.73
Figure of merit	0.75					

† $R_{sym} = \sum |I(h) - \langle I(h)\rangle| / \sum I(h)$, where $\langle I(h)\rangle$ is the mean value of the reflection h for all measurements of $I(h)$. ‡ $R_{anom} = \sum |I^+(h) - I^-(h)| / \sum [I^+(h) + I^-(h)]$, where $I^+(h)$ and $I^-(h)$ are the Bijvoet pairs of $I(h)$. § $R_{iso} = \sum |F_{PH} - F_P| / \sum F_{PH}$, where F_{PH} and F_P are the derivative and native structure-factor amplitudes, respectively. ¶ Phasing power = $\langle F_H\rangle / E$, where $\langle F_H\rangle$ is the mean of the calculated heavy-atom structure-factor amplitude and E is the r.m.s. lack-of-closure error. †† $R_{cullis} = \sum \big||F_{PH} \pm F_P| - F_H(\text{calc})\big| / \sum |F_{PH} \pm F_P|$.

16:0-ACP or 18:1-ACP to the *sn*-1 position of glycerol-3-phosphate (G3P). Within the chloroplast are several thioesterases that compete for the same 16- and 18-carbon ACPs. Hydrolyzed fatty acids are re-esterified to CoA and diverted to the extrachloroplastic pathway of glycerolipid biosynthesis in the endoplasmic reticulum. GPAT also functions in the extrachloroplastic pathway in a manner similar to that observed in the chloroplastic pathway. Experiments using transgenic tobacco plants with GPAT cDNA showed that both the fatty-acid composition and the chilling tolerance of transgenic tobacco plants changed to that of the donor plant species (Murata *et al.*, 1992). This result supports the hypothesis that GPAT controls this lipid distribution and plays a major role in determining a plant's ability to tolerate chilling temperature. We report here the characterization of chimeric enzymes from two genes, *viz.* squash (chilling-sensitive) and spinach (*Spinacea oleracea*) (chilling-tolerant) GPATs, and the interpretation of their kinetic properties based on structural information from stromal GPAT from squash (*Cucurbita moscata*) at 1.55 Å resolution, determined by multiwavelength anomalous dispersion (MAD) with the use of selenomethionyl (SeMet) GPAT. It is proposed that a cavity formed from the region 128–187 determines the substrate specificity of GPAT.

Table 2
Crystallographic refinement.

Values in parentheses refer to the high-resolution shell.

X-ray source	SPring-8 (BL40B2)
Wavelength (Å)	1.00
Resolution (Å)	1.55 (1.61–1.55)
Unique reflections	61882
R_{sym}	0.035 (0.292)
Redundancy	7.1 (6.7)
$I/\sigma(I)$	9.2 (1.5)
Completeness (%)	97.7 (86.6)
Resolution (Å)	15–1.55 (1.61–1.55)
No. of reflections	58534
R_{work}†	0.201 (0.296)
R_{free}‡	0.217 (0.296)
No. of atoms	3247
Protein	2796
Water	390
Sulfate	25
Glycerol	36
Mean B value (Å^2)	26
R.m.s. deviations	
Bond lengths (Å)	0.022
Bond angles (°)	1.862
Ramachandran angles (%)	
Most favoured (%)	93.6
Additional allowed (%)	6.1

† $R_{work} = \sum \big||F_{obs}| - |F_{calc}|\big| / \sum |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. ‡ R_{free} is the same as R_{work} but for a 5% subset of all reflections.

2. Materials and methods

2.1. Purification and crystallization of squash GPAT

Expression and extract preparation of the wild-type and chimeric GPATs consisting of 368 amino acids (not containing the N-terminal pro-sequence that is removed in the mature form) were performed as previously described (Ferri & Toguri, 1997). The cDNA of mature squash GPAT, in which the first three residues (Glu-Pro-Ala) from the amino-terminal were replaced by Met-Ala-Ser, was cloned into the expression vector pET-17b (Novagen) and transformed into

Escherichia coli BL21(DE3)pLysS (Novagen) strain. The expressed squash GPAT was further purified by anion-exchange HPLC using a HiTrap Q column (0.7 × 2.5 cm, Pharmacia). The protein was eluted with a linear gradient from 20 m*M* sodium phosphate buffer pH 8.0 to the same buffer containing 0.5 *M* NaCl. The purified GPAT was concentrated to 14 mg ml^{-1} in 20 m*M* sodium phosphate pH 8.0. Crystals were grown at 277 K in hanging drops over 1.6–1.8 *M* ammonium sulfate over a period of about one week and belong to space group $P2_12_12_1$, with unit-cell parameters a = 60.1, b = 63.7, c = 116.2 Å. These parameters were approximately 13 Å larger along the c axis than those reported by Turnbull and coworkers (a = 61.1, b = 65.1, c = 103.3 Å; Turnbull, Rafferty, Sedelnikova, Slabas, Schierer, Kroon, Nishida *et al.*, 2001). The crystallization conditions [4–25% PEG 4000, 100 m*M* ammonium acetate, 10%(*v*/*v*) 2-propanol in 100 m*M* citrate buffer pH 5.6] of our crystals and their crystal morphology were also different. The Matthews coefficient (V_M) was 2.6 Å^3 Da^{-1} assuming one molecule in the asymmetric unit.

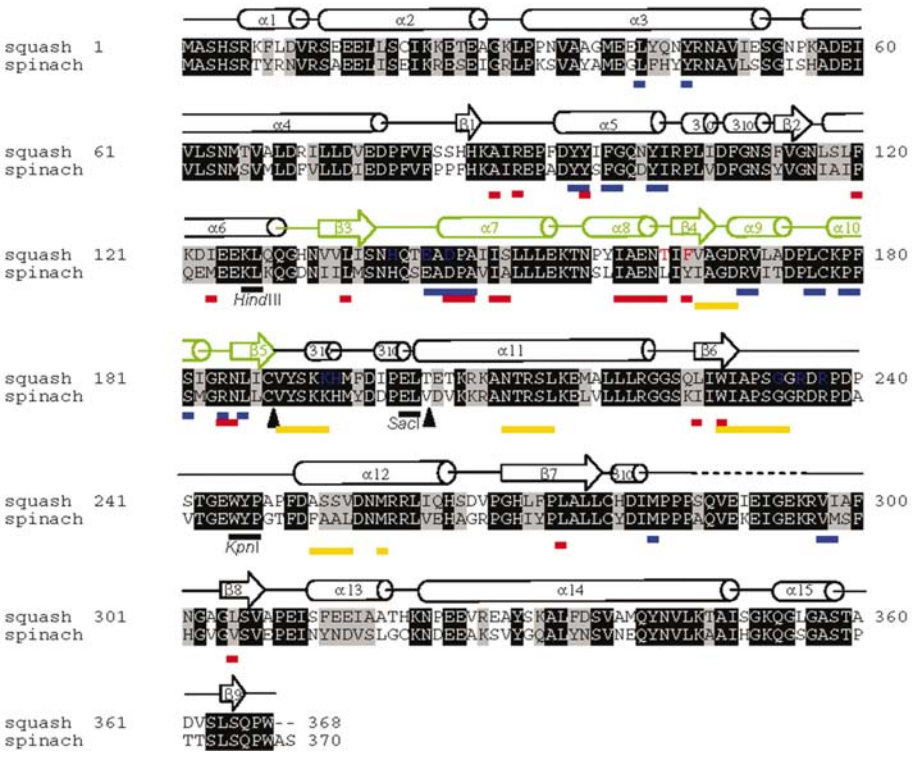

Figure 1
Amino-acid sequence alignment of the squash and spinach GPATs. Alignment was performed using the *CLUSTALX* algorithm (Jeannmougin *et al.*, 1998). The secondary-structure elements (where α is α-helix, 3_{10} is 3_{10}-helix and β is β-strand) derived from the squash GPAT crystal structure are also shown. Identical and similar amino-acid residues are shaded black and grey, respectively. Restriction-enzyme sites (*Hin*dIII, *Sac*I and *Kpn*I) related to the construction of chimeric GPATs are indicated. Two filled arrows located at 188 and 202 represent the *Sac*I sites utilized in the exchange of the small central squash and spinach regions. The essential region for controlling acyl-chain specificity, as determined by the chimeric study, is shown by the secondary-structure elements coloured green. The residues that constitute tunnel-1, tunnel-2 and tunnel-3 are underlined in blue, red and yellow, respectively. In the squash GPAT sequence, the related residues of G3P binding and acyl-chain recognition are shown in blue and red, respectively.

2.2. Preparation of selenomethionyl protein

Selenomethionyl GPAT (SeMet-GPAT) was prepared using a non-auxotrophic strain for methionine (wild strain). The transformant was cultivated in 100 ml of LeMaster medium (LeMaster & Richards, 1985) overnight at 310 K and then diluted with 3 l of LeMaster medium [1/30(*v*/*v*)] containing 50 mg l^{-1} seleno-DL-methionine (Sigma). Purification of SeMet-GPAT was carried out in the same manner as that of wild-type GPAT. Incorporation of selenium into GPAT was confirmed by matrix-assisted laser desorption ionization time-of-flight (MALDI–TOF) mass spectrometry using a Perseptive Biosystem Voyager-DE/RP (Jensen *et al.*, 1996). The observed molecular weight of the GPAT thus obtained was 41 121.7 Da, which is almost equivalent to the theoretical value of 41 116.4 Da obtained by assuming that the seven methionines in squash GPAT were all replaced by selenomethionine (indicating that almost all methionine was replaced by selenomethionine). SeMet-GPAT was crystallized under similar conditions to those used for the wild type and the space group of the SeMet-GPAT crystal was the same as that of the wild type.

2.3. Data collection and structure determination of squash GPAT

MAD measurements were performed at BL18B (Photon Factory) and BL40B2 (SPring-8). Data were collected at 100 K from a single crystal that had been soaked in a solution containing 1.8 *M* ammonium sulfate and 20% glycerol as a cryoprotectant and then flash-frozen in a nitrogen-gas stream. Three data sets were collected on and around the selenium *K* absorption edge at both beamlines using the same crystal. Intensity data from the MAD measurements were integrated using *MOSFLM* (Leslie *et al.*, 1986) and scaled with *SCALA* (Evans, 1993). Absorption-peak data collected at SPring-8 were used for determination of the selenium positions using *SHELX*97 (Sheldrick & Schneider, 1997). Maximum-likelihood heavy-atom refinement and phasing calculations were performed using *MLPHARE* (Otwinowski, 1991), using four selenium positions in conjunction with all six data sets to 2.5 Å resolution. Data-collection and phasing statistics are summarized in Table 1. Phases were improved by solvent flattening and flipping using the programs *DM* (Cowtan, 1994) and *SOLOMON*

(Abrahams & Leslie, 1996). An atomic model was built using the graphics program *QUANTA* (Accelrys Inc., San Diego, CA, USA). Electron-density maps were of sufficient quality for model construction of the entire GPAT molecule, except for residues 286–294.

2.4. Crystallographic refinement of squash GPAT

Full data sets for crystallographic refinement were collected at BL40B2 (SPring-8). The intensity data were processed using *DENZO* and merged with *SCALEPACK* (Otwinowski & Minor, 1997). Crystallographic refinement was carried out by energy minimization and simulated annealing with molecular dynamics using the program *CNX* (Accelrys Inc., San Diego, CA, USA) and by maximum-likelihood refinement using the program *REFMAC* (Murshudov *et al.*, 1997), with manual rebuilding at each refinement cycle. The final model, including 357 residues (Ala2–Pro285 and Arg296–Trp368), 390 water molecules, six glycerol molecules and five sulfate ions, was refined to a crystallographic R factor of 20.1% ($R_{\rm free}$ = 21.9%), using all 58 534 reflections in the resolution range 15.0–1.55 Å (Table 2). The stereochemistry of the refined atomic model, analyzed by *PROCHECK* (Laskowski *et al.*, 1993), showed that all of the main-chain atoms fall within the allowed regions of the Ramachandran plot and the side-chain geometry is within the expected regions at this resolution for the χ_1 and χ_2 stereochemical parameters.

2.5. Preparation of chimeric GPATs

Comparison of the nucleotide sequences between the cDNAs derived from squash and spinach revealed that 70% of the nucleotide sequence was identical (Fig. 1). Several restriction sites were found in the conserved sequences and *Hin*dIII and *Kpn*I cleavage sites were present at positions that divide the enzyme almost equally into three portions (Fig. 2*a*). Using these sites, the DNA was recombined to create six different GPAT chimeras (Fig. 2*a*; Ferri & Toguri, 1997). The resulting chimeras were given three-letter names, incorporating the second letters of 'S**P**INACH' and 'S**Q**UASH' to indicate the origins of each segment of the GPAT chimera. In addition, a finer chimera for the central part was prepared utilizing a *Sac*I site within the central portion, Q, of PQP; these chimeras were given four or five-letter names with parentheses (Fig. 2*a*). A synthetic DNA, P1 (5′-AC**GAGCTC**GGGATC-ATCATACATGTGCTT-3′), where the *Sac*I site is in bold, was designed in the antisense fashion from the site corresponding to the centre region of the spinach gene. A synthetic DNA (5′-TGACGCATGCGCTAGCCACTCTCGCACTTATCGTAA-CGTTCGT-3′), which had been used for the construction of previous chimeric genes with spinach-derived amino-termini, and P1 were combined and PCR was carried out under the conditions described by Ferri & Toguri (1997). The DNA fragment thus produced was cut with *Hin*dIII and *Sac*I to give a fragment of the central part of the gene derived from spinach which can be introduced between the *Sac*I and *Hin*dIII sites of the squash gene. After PCR with the squash gene as a template in the manner described above, the DNA fragment was cut with *Hin*dIII and *Sac*I and the fragment that had been inserted between the recognition sites of these enzymes was obtained and recovered. These DNA fragments and a plasmid in which PQP had been cut with *Hin*dIII and *Sac*I and the fragment between them removed were combined to prepare P(PQ)P and P(QPQ)P, respectively. In these constructions, the sequence derived from spinach that is present in the centre of the latter chimeric gene is the sequence corresponding to the primer used in PCR. Also, the central fragment (PQ) of the chimera obtained by cutting P(PQ)P with the restriction enzymes *Hin*dIII and *Kpn*I was inserted in place of P, which had been removed by digestion of QPQ with the same set of enzymes to form Q(PQ)Q. Further construction, expression and preparation were carried out in the same manner as for the wild type (Ferri & Toguri, 1997). Nine different chimeras were ultimately created (Fig. 2*a*).

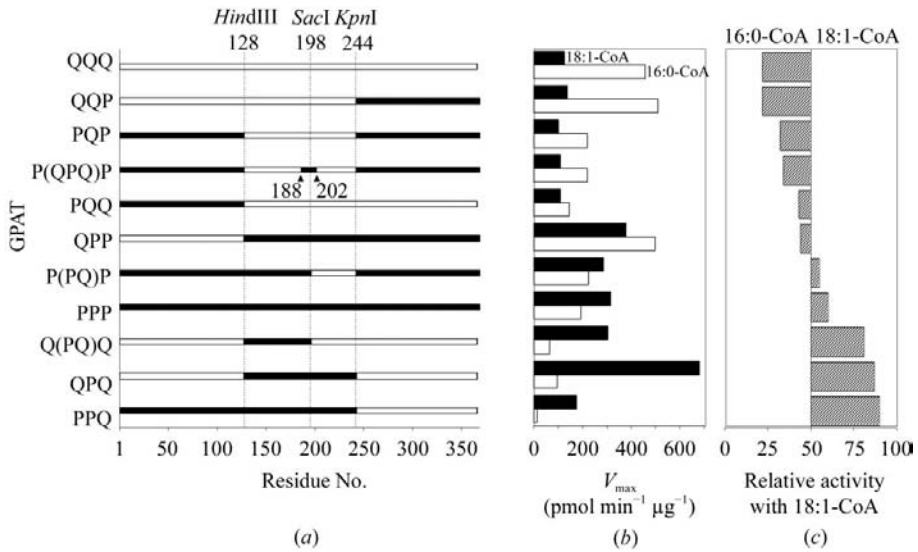

Figure 2
Construction, enzymatic activities and substrate specificities of chimera GPATs. (*a*) Construction of chimeric GPATs. Three-letter names composed of the second letters of 'S**Q**UASH' (colored white) and 'S**P**INACH' (black) are used to indicate the origins of each segment of the GPAT chimera. The four- or five-letter names with parentheses represent the chimeras further subdivided by exchanging sequences using the *Sac*I sites. (*b*) Enzymatic activity of chimeric GPATs towards 16:0-CoA and 18:1-CoA. Enzymatic activity was expressed as the maximum transfer rate ($V_{\rm max}$) of the acyl chain from acyl-CoA to (U-^{14}C)-G3P. (*c*) Relative activities expressed as a ratio of $V_{\rm max}$ for 16:0-CoA to the sum of $V_{\rm max}$ for 18:1-CoA and $V_{\rm max}$ for 16:0-CoA (data taken from Fig. 2*b*).

2.6. Enzymatic analysis of the *C. moscata* (squash), *S. oleracea* (spinach) and chimeric GPATs

The enzymatic activities of the wild-type and chimeric GPATs were determined by measuring the transfer rate

of the acyl chain from acyl-CoA to (U-^{14}C)-G3P, as described previously (Bertrams & Heinz, 1981), with minor modifications. Approximately 1 µg of crude protein was incubated at 297 K in 80 µl 0.25 *M* HEPES pH 7.4 containing 6 µg ml^{-1} bovine serum albumin, 0.3 m*M* (U-^{14}C)-G3P (0.9 Ci mol^{-1}) and 0.4 m*M* 16:0- or 18:1-CoA for 8 min and the reaction was then stopped by adding 2.3 ml of a 1:1 mixture of chloroform and methanol and 1 ml each of 1 *M* KCl and 0.2 *M* H_3PO_4. After a quick centrifugation, 0.9 ml of the lower organic layer was mixed with scintillation cocktail (Aquasol-2) and counted. The amount of protein extract added was varied to maintain the product concentration below 30% of the maximum. All assays were performed in triplicate and the standard deviation was confirmed to be less than 2% of the values reported previously (Ferri & Toguri, 1997).

3. Results

3.1. Enzymatic activities of chimeric GPATs

In order to elucidate the regions responsible for the substrate selectivity of GPAT, several chimeric GPATs between squash and spinach were constructed (Fig. 2*a*). The enzymatic activities of these chimeric GPATs were measured in triplicate (within 2% standard deviation), as shown in Fig. 2(*b*). It was found that none of the chimeric GPATs were inactivated (Fig. 2*b*). The enzymatic activities of chimeric GPATs, however, clearly depend on the substrate. The wild-type squash GPAT (QQQ) prefers 16:0-CoA, while spinach GPAT (PPP) prefers 18:1-CoA. The chimera having the highest enzymatic activity for 18:1-CoA was QPQ, whereas that showing highest enzymatic activity for 16:0-CoA was QQP.

In Fig. 2(*c*), the relative activities of the wild-type and chimeric GPATs are shown in order of substrate selectivity. The relative selectivities were expressed as a ratio of V_{max} for 16:0-CoA to the sum of V_{max} for 18:1-CoA and V_{max} for 16:0-CoA. It was found that the wild-type squash GPAT (QQQ) exhibited higher selectivity for 16:0-CoA than 18:1-CoA and that the wild-type spinach GPAT (PPP) prefers 16:0-CoA to 18:1-CoA.

The enzymatic activity of these chimeric GPATs revealed an important region for substrate selectivity. The chimeric GPATs having the spinach sequence from residues 128–198 [QPQ, PPP, Q(PQ)Q, P(PQ)P and PPQ] showed higher specificity for 18:1-CoA than for 16:0-CoA, suggesting that this spinach sequence is important for the recognition of 18:1-CoA. The relative activities of PPQ, QPQ and Q(PQ)Q towards 18:1-CoA were higher than that of the wild-type spinach GPAT (PPP). It is interesting that QPQ showed the highest enzymatic activity (more than twice as active as PPP) of these chimeric GPATs. It was concluded that a combination of the region derived from spinach in the central part and the region from squash in the C-terminal part was optimal for selectivity. Only QPP of the chimeras with the spinach sequence in the central part displayed a reversed relative activity to that of the wild-type enzyme, although QPP also showed significant activity towards 18:1-CoA. On the other hand, the GPATs having the sequence 128–244 (except 188–202) derived from squash [QQP, QQQ, PQP, P(QPQ)P, PQQ] generally show higher activities towards 16:0-CoA, so this sequence was found to be important for reactions using

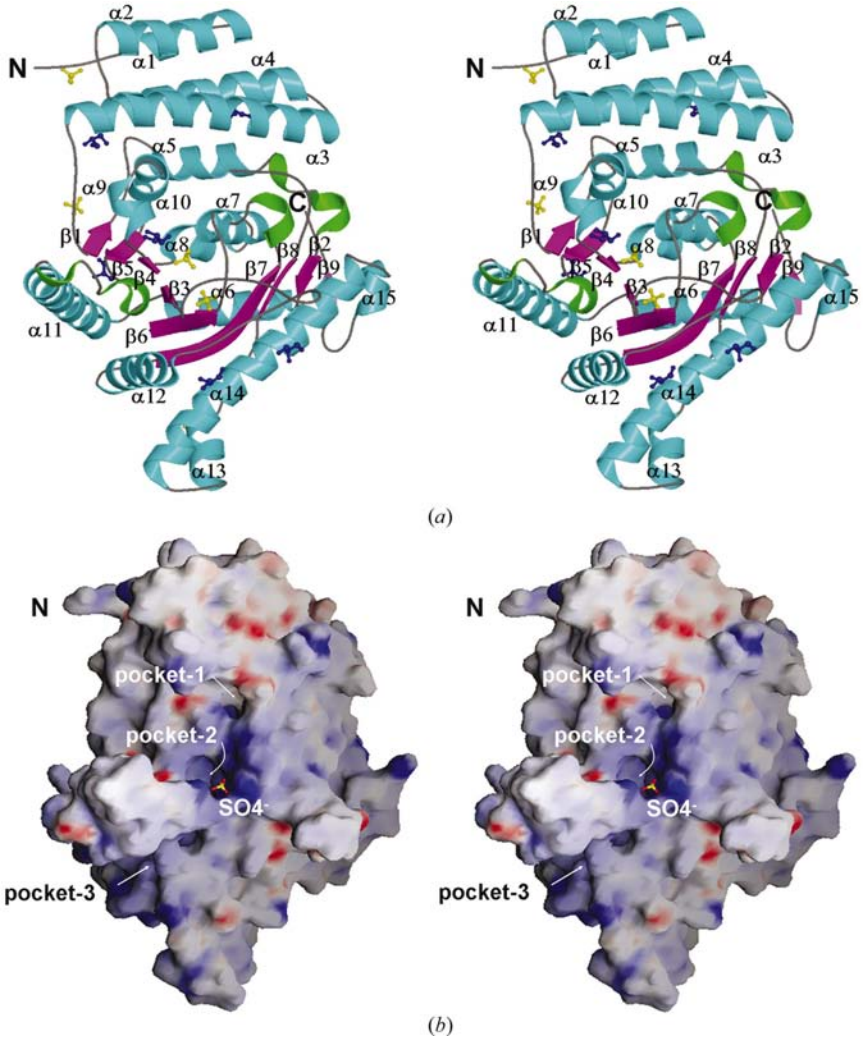

Figure 3
Tertiary structure of squash GPAT determined by X-ray crystallography. (*a*) Ribbon diagram of the overall structure of squash GPAT. The β-strands are colored magenta, α-helices cyan and 3_{10}-helices green. The glycerols and sulfate ions are drawn in ball-and-stick representations in blue and yellow, respectively. This figure was prepared with *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Murphy, 1994). (*b*) Electrostatic surface potential of squash GPAT was calculated and displayed with *GRASP* (Nicholls *et al.*, 1991). Blue and red indicate positive and negative charged residues, respectively. A sulfate ion in pocket-2 is drawn as a stick model.

saturated fatty acids (16:0-CoA). From these results of the chimeric study, the 60 amino-acid residues 128–187 are responsible for controlling acyl-chain specificity.

In the comparison of the overall V_{max}, it is evident that those constructs with the amino-terminal region (1–128) from squash exhibit generally higher V_{max} values than those with the similar region from spinach. Thus, this region may have an overall effect upon enzymatic activity that is independent of selectivity.

3.2. Crystal structure of GPATs from *C. moscata* (squash)

The crystal structure of *C. moscata* (squash) GPAT was determined at 1.55 Å resolution by MAD phasing using a selenium-containing protein crystal. The polypeptide chain of squash GPAT is folded into nine β-strands, 15 α-helices, five short 3_{10}-helices and connecting segments (Fig. 3*a*). The structure is composed of two domains: a helical domain and an α/β domain. The approximate overall dimensions are 70 × 50 × 50 Å. Residues 1–77 form the helical domain and the α/β domain comprises residues 85–368. A short loop, residues 78–84, connects the two domains. The helical domain has four helices (α1–α4) which constitute a four-helix bundle. This domain resembles a slab with approximate dimensions 40 × 25 × 20 Å. The α/β domain is folded into 11 α-helices (α5–α15), five 3_{10}-helices and nine β-strands (β1–β9). The overall size of the α/β domain is 70 × 50 × 50 Å. The core of the α/β domain has an open twisted α/β structure, which is folded into five parallel β-strands (β3–β7) and three antiparallel β-strands (β1, β2 and β8). These structural features are similar to those of the GPAT structure determined at 1.9 Å resolution by Turnbull and coworkers (Turnbull, Rafferty, Sedelnikova, Slabas, Schierer, Kroon, Simon *et al.*, 2001; PDB code 1k30), except for the presence of five glycerol molecules used as cryoprotectant and four sulfate ions used as precipitant in crystallization around our protein structure. The r.m.s.d. of the overall structures of these GPATs calculated using C^{α} atoms is 0.91 Å and the r.m.s.d. values of the helical and α/β domains calculated using C^{α} atoms are 0.73 and 0.83 Å, respectively.

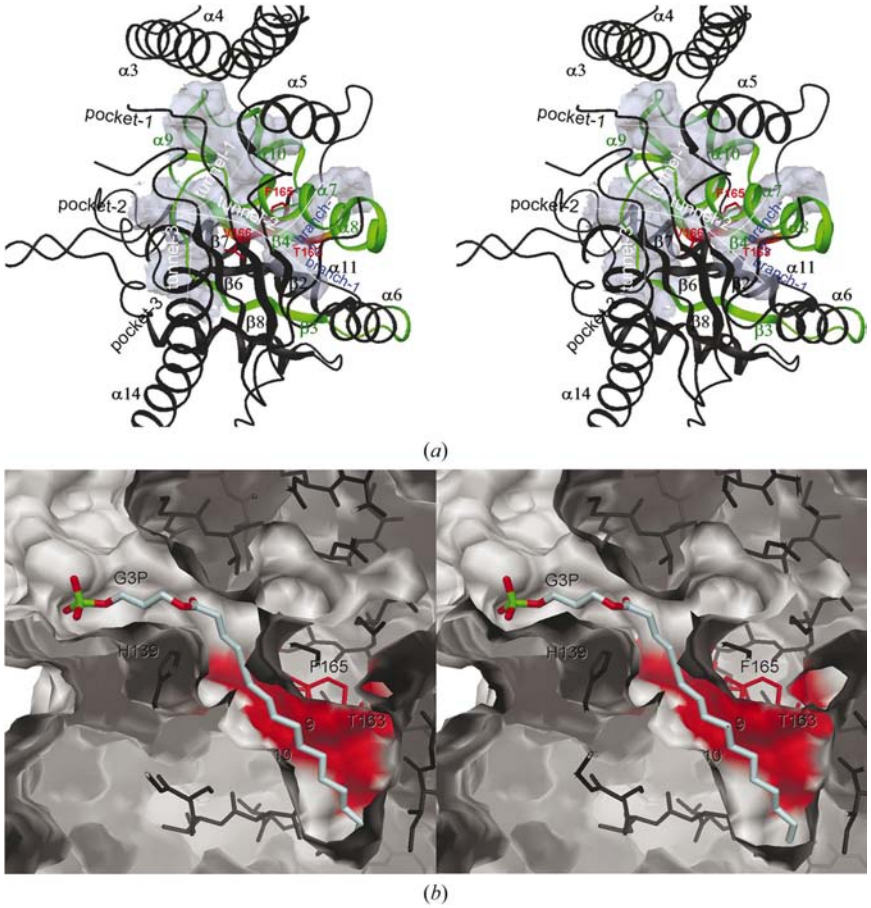

Figure 4
The locations of pockets and tunnels seen in the structure of squash GPAT. (*a*) Relative locations of pocket-1–pocket-3 and tunnel-1–tunnel-3. The locations of tunnel-1 and tunnel-3 are shown as dotted lines and the locations of tunnel-2 and its branches are shown as white lines. The essential region (residues 128–187) for controlling acyl-chain specificity, as determined by the chimeric study, is coloured green. The side chains of the three non-identical residues, Thr163, Phe165 and Val166, that construct tunnels in this essential region are shown in red. Cavities were calculated with the program *VOIDOO* (Kleywegt & Jones, 1994). (*b*) Putative location of 16:0-G3P in the structure of GPAT. Two residues, Thr163 and Phe165, are located near the unsaturated bond between the ninth and tenth C atom of 18:0-G3P. The model was built by 5 ps simulation of molecular-dynamics calculations at 300 K using the program *DISCOVER* (Accelrys Inc., San Diego, CA, USA). The initial model was produced using the software *AFFINITY* (Accelrys Inc., San Diego, CA, USA) by considering the results of chimeric and structural studies. The molecular surface was calculated using the program *MSMS* (Sanner *et al.*, 1996). Figures were drawn using the program *DINO* (http://www.dino3d.org).

Moreover, molecular-surface calculations using 1.1 Å diameter as a probe radius revealed that there are obvious cavity features, including three pockets (pocket-1–pocket-3) and three tunnels (tunnel-1–tunnel-3), in the tertiary structure of stromal GPAT (Figs. 3*b* and 4*a*). The pockets are ∼6 Å in diameter (Fig. 3*b*) and tunnel-1 and tunnel-3 extend from pocket-2 in the directions of pocket-1 and pocket-3, respectively. Tunnel-2 is hydrophobic, extending from pocket-2 perpendicular to the other tunnels, and is further divided into branch-1 and branch-2 in the middle of tunnel-2, which was not mentioned in previous reports (Turnbull, Rafferty, Sedelnikova, Slabas, Schierer, Kroon, Simon *et al.*, 2001). Branch-1 did not reach the molecular surface, but branch-2 reached the surface of stromal GPAT. Each tunnel appears to have the necessary dimensions to contain the acyl chain of 16:0-CoA (18.5 Å). The r.m.s.d. values between the structures reported here and that of Turnbull and coworkers for all atoms constituting tunnel-1 (the 25 residues shown in blue in Fig. 1), tunnel-2 (the 23 residues in red) and tunnel-3 (the 26 residues in yellow) were calculated to be 1.60, 1.69 and 1.32 Å, respectively.

Pocket-2 contains the positively charged residues His139, Lys193, His194, Arg235 and Arg237 seen in the blue cluster in Fig. 3(*b*). A strong electron density interacting with these cationic residues (Fig. 5*a*) was found at pocket-2. The density is located within direct hydrogen-bonding distance of the main-chain and side-chain N atoms of Arg235, the side-chain N atoms of Lys193, His194 and Arg237, the main-chain O atom of Gly233 and the side-chain N atom of His139 (mediated by a water molecule). It was also noticed that His139, Lys193 (corresponding to His306 and Arg354 in *E. coli* GPAT) and Gly233 (corresponding to Gly386 in *E. coli* GPAT) were important for catalysis of *E. coli* GPAT (Lewin *et al.*, 1999) and that Arg235 (corresponding to Arg318 in murine mitochondrial GPAT) was important for murine mitochondrial GPAT activity (Dircks *et al.*, 1999). The clear electron density found in pocket-2 is interpreted as a sulfate ion used as precipitant in crystallization. Since sulfate ions were observed at the G3P-binding site surrounded by several positively charged residues, this position may be a G3P-binding site of stromal GPAT. Although co-crystallization with substrates (16:0-CoA or 18:1-CoA and G3P) and their analogues was performed to investigate the substrate-binding site of GPAT, no electron density from substrates could be observed even with a tenfold molar excess of each substrate and crystals appeared to contain only the apo-form of the enzyme.

4. Discussion

Selective incorporation of saturated (16:0) and unsaturated (18:1) fatty-acyl chains into chloroplast membranes by stromal GPATs is a key determinant of a plant's ability to tolerate chilling temperatures. Therefore, it was deemed important to elucidate the mechanism of substrate recognition by squash (chilling-sensitive) and spinach (chilling-tolerant) GPATs. Although tertiary-structure information on the enzyme–substrate complex is not yet available for any GPATs, enzymatic analysis of chimeric proteins between squash (chilling sensitive) and spinach, as well as high-resolution X-ray crystallography at 1.55 Å of stromal GPAT, enables us to understand the molecular basis of the catalytic function of stromal GPAT. The enzymatic analysis of the chimeric GPATs suggests that the region 128–187 is responsible for fatty-acid recognition. The tertiary structure of GPAT showed two key features, the existence of a basic residue cluster and the hydrophobic tunnel in the structure of stromal GPAT, which may be useful for predicting the location of the substrate-binding site.

One of the key structural features is the existence of a basic residue cluster consisting of His139, Lys193, His194, Arg235 and Arg237 that interacts with a strong electron density belonging to a sulfate ion within pocket-2. This cluster had been predicted to be a G3P-binding site by Turnbull and coworkers (Turnbull, Rafferty, Sedelnikova, Slabas, Schierer, Kroon, Simon *et al.*, 2001) and it is proposed that His139 and the adjacent Asp144 in this cluster promote a charge-relay system in squash GPAT (from the sequence similarity to His306 and Asp311 in *E. coli* GPAT). His306 is known to act as a general base to deprotonate the hydroxyl moiety of the G3P in order to facilitate nucleophilic attack on the thioester of the acyl-CoA with the assistance of Asp311 in *E. coli* GPAT (Heath & Rock, 1998). The contributions of the charge-relay system to the transfer reaction of the acyl chain from CoA to an enzyme hydroxyl group is also confirmed from the crystal structure determination of the palmitoyl protein thioesterase 1 (PPT1)–palmitate (16:0) complex (Bellizzi *et al.*, 2000), in which the hydroxyl group of a serine residue acts as an acyl acceptor instead of that of G3P in *E. coli* GPAT. Therefore, it is suggested that pocket-2 is the G3P-binding site of stromal GPAT.

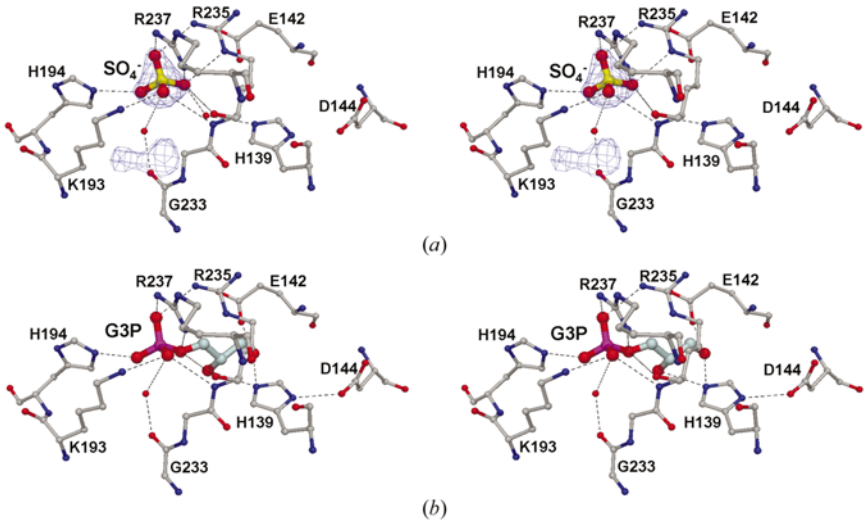

Figure 5
Positively charged cluster and the interaction with (*a*) the sulfate ion and (*b*) G3P. The $F_o - F_c$ omit electron-density maps contoured at 4σ corresponding to the sulfate ion are also shown (in cyan). Dashed lines indicate hydrogen bonds. The putative binding model of G3P was obtained using the program *DISCOVER* (Accelrys Inc., San Diego, CA, USA). Figures were drawn using the program *DINO* (http://www.dino3d.org).

The location of the G3P-binding site in stromal GPAT in pocket-2 is also supported by the strong electron density from the sulfate ion. It is known that sulfate ions sometimes occupy phosphate-binding sites in crystal structures. Indeed, a sulfate ion was observed at the same position as a phosphate group of 2-phosphoglycolate (2PG), which was a substrate analogue, in holocrystals of apo-triose-phosphate isomerase (TIM; Alvarez *et al.*, 1998). In order to estimate the probability of pocket-2 being the binding site for G3P, a model structure was built by replacing the sulfate ion with G3P. In the model structure, the hydrogen-bonding network around the sulfate ion was conserved around the phosphate group of G3P (Fig. 5*b*). The *sn*-1 hydroxyl group of G3P converges to a position in which it is possible for it to interact with the side-chain $N^{\varepsilon 2}$ atom of His139 and the side-chain $O^{\varepsilon 2}$ atom of Glu142. The side-chain $O^{\delta 2}$ atom of Asp144

is within direct hydrogen-bonding distance of the $N^{\delta1}$ atom of His139. It seems that this model can explain the fact that the Asp-His charge relay increases the nucleophilicity of the *sn*-1 hydroxyl group of G3P, which attacks the acyl-CoA to form acyl-G3P.

Another key feature observed in the crystal structure of squash GPAT was the existence of the three long tunnels (tunnel-1–tunnel-3) located deep inside squash GPAT, as shown in Fig. 4(*a*). Every tunnel appears to have sufficient space to accommodate the acyl chain of 16:0-CoA. Tunnel-2 is the narrowest and most hydrophobic of the three tunnels. From the structural point of view, the existence of such a hydrophobic tunnel should result in significant destabilization of the enzyme and it is therefore reasonable to consider that tunnel-2 must be involved in some function of GPAT. If G3P binds to pocket-2, there are three possibilities for acyl-chain binding, namely the locations between pocket-2 and each of the three tunnels (see Fig. 4*a*). These possibilities can be examined according to the enzymatic analysis of chimeric GPATs (Fig. 2), as well as the structural characteristics of the tunnels. As mentioned previously, the region containing residues 128–187 has been identified as an important region for distinguishing the selectivity between squash and spinach GPATs (Figs. 2*b* and 2*c*). It was found that tunnel-1 is partially covered and tunnel-2 is completely covered by this region (shown in green in Figs. 1 and 4*a*), whereas tunnel-3 is hardly covered by this region. It has already been predicted from molecular modelling of the GPAT–acyl chain complex that tunnel-3 is the binding site of the acyl chain (Turnbull, Rafferty, Sedelnikova, Slabas, Schierer, Kroon, Simon *et al.*, 2001). The choice of tunnel-3, however, seems to conflict with the results of our chimeric study and tunnel-3 may not be suitable for the binding site of the acyl chain as it is more hydrophilic than the other tunnels. This hypothesis is also supported by the existence of several water molecules.

17 residues from the region 128–187 are not identical in the squash and spinach sequences (Fig. 1). Of these non-identical residues, only three residues (Thr163, Phe165 and Val166) participate in the construction of tunnels. These three residues (shown in red in Fig. 4) are not located in tunnel-1 but are located in the vicinity of tunnel-2, suggesting that tunnel-2 is the acyl-chain binding site responsible for controlling acyl-chain specificity. Moreover, the acyl chain of acyl-G3P was fitted into tunnel-2 by molecular modelling based on the information of the location of G3P (Fig. 4*b*) and subjected to a molecular-dynamics simulation. Only small positional shifts (r.m.s.d. values between the X-ray and model structures calculated using C^{α} atoms were 0.07 Å for all residues and 0.15 Å for the region 127–187) were observed, suggesting reasonable fitting of the acyl chain into tunnel-2 of squash GPAT. The depth of branch-1 in tunnel-2 is an appropriate size for docking of 16:0-G3P, whereas 18:1-G3P cannot fit into branch-1. The branch point showed a similar location for the unsaturated bond of 18:0-G3P (between the ninth and tenth C atoms). The structural characteristics of tunnel-2 in squash GPAT suggest that it possesses higher specificity for 16:0-CoA. Since two non-identical residues between squash and spinach GPAT, Thr163 (Leu163 in spinach) and Phe165 (Tyr165 in spinach), are located at this branch point, these residues may play a key role in control of specificity of stromal GPATs. Although detailed analysis of the identification of the acyl-chain binding site is needed to fully understand how stromal GPATs distinguish saturated and unsaturated fatty-acyl chains, we propose here that tunnel-2 is the acyl-chain binding site. We also speculate that the acyl-CoA may initially bind to either of the adjacent tunnels (1 or 3) as a non-productive binding event and an associated structural perturbation may allow the hydrophobic acyl chain to penetrate into the hydrophobic region leading to tunnel-2. These findings will help us to understand the molecular mechanism of the chilling tolerance of plants.

We thank Drs Keiko Miura and Masahide Kawamoto of SPring-8 (proposal Nos. 2000B0288-NL-np, 2001A0481-CL-np and 2001B0254-NL-np), and Drs Noriyoshi Sakabe and Noriyuki Igarashi of the Photon Factory for data collection with synchrotron radiation. TT and RK are members of the SBSP (Structural Biology Sakabe Project) of the FAIS (Foundational Academy of International Science). We are also indebted to Dr Haruhiko Tamaoki of Kumamoto University School of Medicine for supplying the substrate analogues, Azusa Kanda-Shinohara for MALDI–TOF MS measurements and Vicki Leung for preparation of SeMet-GPAT.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Alvarez, M., Zeelen, J. P., Mainfroid, V., Reintier-Delrue, F., Martial, J. A., Wyns, L., Wierenga, R. K. & Maes, D. (1998). *J. Biol. Chem.* **273**, 2199–2206.
Bellizzi, J. J. III, Widom, J., Kemp, C., Lu, J.-Y., Das, A. K., Hofman, S. L. & Clardy, J. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 4573–4578.
Bertrams, M. & Heinz, E. (1981). *Plant Physiol.* **68**, 653–657.
Bishop, D. G. (1986). *Plant Cell Environ.* **9**, 613–616.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
Dircks, L. K., Ke, J. & Sul, H. S. (1999). *J. Biol. Chem.* **274**, 34728–34734.
Evans, P. R. (1993). In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey. Warrington: Daresbury Laboratory.
Ferri, S. R. & Toguri, T. (1997). *Arch. Biochem. Biophys.* **337**, 202–208.
Frentzen, M. (1993). *Lipid Metabolism in Plants*, edited by T. S. Moore Jr, pp. 195–231. Boca Raton, FL, USA: CRC Press.
Harwood, J. L. (1996). *Biochim. Biophys. Acta*, **1301**, 7–56.
Heath, R. J. & Rock, C. O. (1998). *J. Bacteriol.* **180**, 1425–1430.
Jeannmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998). *Trends Biochem. Sci.* **23**, 403–405.
Jensen, O. N., Podtelejnikov, A. & Mann, M. (1996). *Rapid Commun. Mass Spectrom.* **10**, 1371–1378.
Kleywegt, G. J. & Jones, A. T. (1994). *Acta Cryst.* D**50**, 178–185.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacAuthor, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
LeMaster, D. M. & Richards, F. M. (1985). *Biochemistry*, **24**, 7263–7268.

Leslie, A. G. W., Brick, P. & Wonacott, A. T. (1986). *CCP4 Newsl.* **18**, 33–39.
Lewin, T. M., Wang, P. & Coleman, R. A. (1999). *Biochemistry*, **38**, 5764–5771.
Lyons, J. M. (1973). *Annu. Rev. Plant Physiol.* **24**, 445–466.
Merritt, E. A. & Murphy, M. E. (1994). *Acta Cryst.* D**50**, 869–873.
Murata, N. (1983). *Plant Cell Physiol.* **24**, 81–86.
Murata, N., Ishizaki-Nishizawa, O., Higashi, S., Hayashi, H., Tasaka, Y. & Nishida, I. (1992). *Nature* (*London*), **356**, 710–713.
Murata, N., Sato, N., Takahashi, N. & Hamazaki, Y. (1982). *Plant Cell Physiol.* **23**, 1071–1079.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Ohlrogge, J. & Browse, J. (1995). *Plant Cell*, **7**, 957–990.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Roughan, P. G. (1985). *Plant Physiol.* **77**, 740–746.
Sanner, M. F., Olson, A. J. & Spehner, J.-C. (1996). *Biopolymers*, **38**, 305–320.
Sheldrick, G. & Schneider, T. (1997). *Methods Enzymol.* **277**, 319–344.
Turnbull, A. P., Rafferty, J. B., Sedelnikova, S. E., Slabas, A. R., Schierer, T. P., Kroon, J. T. M., Nishida, I., Murata, N., Simon, J. W. & Rice, D. W. (2001). *Acta Cryst.* D**57**, 451–453.
Turnbull, A. P., Rafferty, J. B., Sedelnikova, S. E., Slabas, A. R., Schierer, T. P., Kroon, J. T. M., Simon, J. W., Fawcett, T., Nishida, I., Murata, N. & Rice, D. W. (2001). *Structure*, **9**, 347–353.