Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Binding of inhibitory aromatic amino acids to *Streptomyces griseus* aminopeptidase

V. Reiland,[a] R. Gilboa,[a] A. Spungin-Bialik,[b] D. Schomburg,[c] Y. Shoham,[d] S. Blumberg[b] and G. Shoham[a]*

[a]Department of Inorganic Chemistry and the Laboratory for Structural Chemistry and Biology, The Hebrew University of Jerusalem, Jerusalem 91904, Israel, [b]Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel, [c]Institute of Biochemistry, University of Cologne, Cologne 50674, Germany, and [d]Department of Biotechnology and Food Engineering and Institute of Catalysis Science and Technology, Technion–Israel Institute of Technology, Haifa 32000, Israel

Correspondence e-mail: gil2@vms.huji.ac.il

Received 10 June 2004
Accepted 26 July 2004

PDB References:
SGAP–L-Trp, 1tf8, r1tf8sf;
SGAP–*p*-I-L-Phe, 1tf9, r1tf9sf.

The bacterial aminopeptidase isolated from the extracellular extract of *Streptomyces griseus* (SGAP) is a double-zinc exopeptidase with a high preference for large hydrophobic amino-terminus residues. It is a monomer of a relatively low molecular weight (30 kDa), is heat-stable, displays a high and efficient catalytic turnover and its activity is modulated by calcium ions. Several free amino acids were found to inhibit the activity of SGAP in the millimolar concentration range and can therefore serve for the study of binding of both inhibitors and reaction products. The current study is focused on the X-ray crystallographic analysis of the SGAP complexes with L-tryptophan and *p*-iodo-L-phenylalanine, both at 1.30 Å resolution. These two bulky inhibitory amino acids were found to bind to the active site of SGAP in very similar positions and orientations. Both of them bind to the two active-site zinc ions *via* their free carboxylate group, while displacing the zinc-bound water/hydroxide that is present in the native enzyme. Further stabilization of the binding of the amino-acid carboxylate group is achieved by its relatively strong interactions with the hydroxyl group of Tyr246 and the carboxylate group of Glu131. The binding is also stabilized by three specific hydrogen bonds between the amine group of the bound amino acid and enzyme residues Glu131, Asp160 and Arg202. These consistent interactions confirm the key role of these residues in the specific binding of the free amine of substrates and products, as proposed previously. The phenyl ring of Phe219 of the enzyme is involved in stacking interactions with the corresponding aromatic ring of the bound affector. This interaction seems to be important for the binding and orientation of the aromatic side chain within the specificity pocket. These structural results correlate well with the results obtained for the complexes of SGAP with other inhibitory amino acids and support the general catalytic mechanism proposed for this and related enzymes.

1. Introduction

Aminopeptidases catalyze the cleavage of amino acids from the amino-terminus of many proteins and peptides. They appear to be widely distributed in bacteria, plants and mammalian tissues and have been found in the cytoplasm and in subcellular organelles (Taylor, 1993*a*,*b*, 1996). Aminopeptidases have been found to be essential for many central biological processes such as protein maturation, hormone-level regulation, cell-cycle control, protein degradation and viral infection (Taylor, 1993*a*,*b*, 1996). When these enzymes are absent, malfunctioning or slowed down, various abnormal conditions and diseases can develop, such as inflammation, cancer, aging, leukaemia, cystic fibrosis and cataracts. A detailed understanding of the structure and catalytic mode of

action of aminopeptidases and especially of the factors that are critical for their proper biological function thus seems to be important for a wide range of medical and pharmacological purposes.

Streptomyces griseus aminopeptidase (SGAP) is an aminopeptidase that shows a high preference for large hydrophobic amino-acid residues at the amino-terminus of peptides and proteins. In common with many aminopeptidases, SGAP contains two zinc ions per protein molecule, both of which seem to be involved in substrate binding and catalysis. In addition to zinc, calcium ions have been shown to affect the activity, specificity and selectivity of SGAP (Spungin & Blumberg, 1989; Ben-Meir *et al.*, 1993; Papir *et al.*, 1998). The original amino-acid sequence of SGAP was determined by direct chemical methods using the native enzyme from the *S. griseus* bacterium (Maras *et al.*, 1996). The enzyme has recently been cloned, overexpressed and purified from *Escherichia coli* (Fundoiano-Hershcovitz *et al.*, 2004). It is a monomer of about 30 kDa, is heat-stable and displays a relatively high and efficient catalytic turnover. These properties make SGAP an excellent representative enzyme for the general study of aminopeptidases. They also make SGAP a very attractive enzyme for a wide range of biotechnological applications, including the processing of recombinant DNA proteins and diagnostic procedures for the activity of other peptidases (Ben-Bassat *et al.*, 1987; Nagakawa *et al.*, 1987; Indig *et al.*, 1989, 1990; Ben-Meir & Blumberg, 1991).

SGAP belongs to a subgroup of aminopeptidases which are not very specific and which are capable of cleaving a number of hydrophobic amino acids with reasonable catalytic rates. In addition to SGAP, this subgroup includes the leucine aminopeptidase from bovine lens (blLAP; Burley *et al.*, 1992) and the aminopeptidase from *Aeromonas proteolytica* (AAP; Chevrier *et al.*, 1994). While LAP appears to be very different from SGAP in molecular weight, sequence and structure, AAP has been shown to be very similar to SGAP in both size, amino-acid sequence (about 30% homology), three-dimensional structure and composition of the active site (Greenblatt *et al.*, 1997; Gilboa *et al.*, 2000). Two other aminopeptidases that have been investigated extensively, methionine aminopeptidase (MAP; Roderick & Matthews, 1993) and proline aminopeptidase (PAP; Taylor, 1996; Wilce *et al.*, 1998), have been shown to have no resemblance to SGAP or AAP at either the sequence or the structure level. Interestingly, several other proteins have recently been shown to have significant sequence homology to SGAP, including the medically important prostate-specific membrane antigen (PSMA; Mahadevan & Saldanha, 1999) and the transferrin receptor (TfR; Lawrence *et al.*, 1999).

The current research is a part of a detailed study of the interactions of SGAP with a wide range of inhibitors and analogues. This study is driven by the strong mechanistic interest in this family of double-zinc aminopeptidases and recent interest in the medical and pharmacological aspects of SGAP and related proteins. To date, the crystal structure of native SGAP has been determined at 1.75 Å (Greenblatt *et al.*, 1997; Papir *et al.*, 1998), 1.58 Å (Gilboa *et al.*, 2000) and 1.35 Å resolution (Gilboa *et al.*, 2004), followed by the high-resolution structure determination of a number of SGAP complexes with various inhibitors (Gilboa *et al.*, 2000, 2001, 2004). These studies showed that SGAP consists of a central β-sheet, which is made up of eight parallel and antiparallel strands and is surrounded by helices on either side (see Fig. 1 below). The active site, which contains the two zinc ions, was shown to be located at the end of a middle strand of the β-sheet region. The calcium-binding site was shown to be located near the amino-terminus of the enzyme, about 25 Å away from the active site. More recently, the structural study of SGAP has been focused on the structural analysis of complexes of the enzyme with a series of amino acids, which serve as products (and product analogues) and at the same time show considerable inhibition of the enzyme. The first three amino acids to be investigated in this way were methionine, leucine and phenylalanine (Gilboa *et al.*, 2000, 2001), all of which act as weak inhibitors of SGAP (K_i values in the range 8–13 $\times$ 10^{-3} M in the presence of Ca^{2+}; Papir *et al.*, 1998).

The present study extends the analysis of SGAP interactions with amino acids of this kind and concentrates on those with bulky aromatic side chains. Here, we present the high-resolution crystal structures of the SGAP complexes with

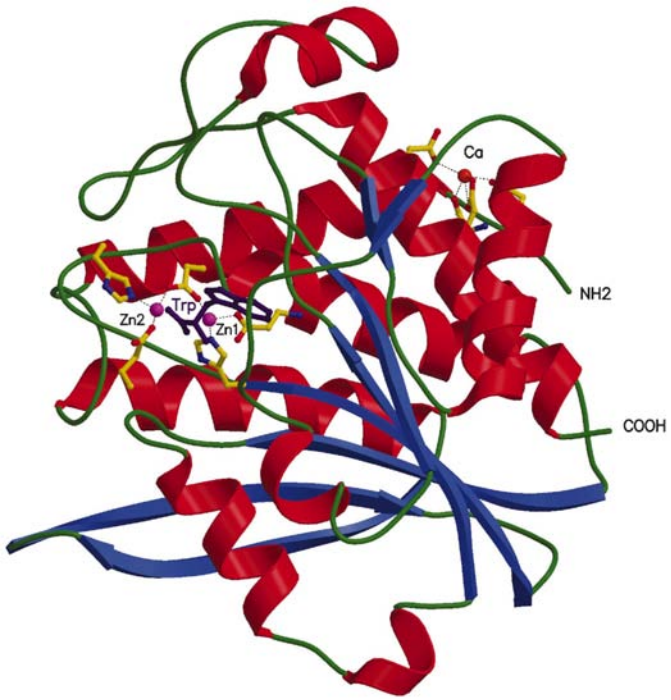

Figure 1
A schematic plot of the overall structure of the SGAP–L-Trp complex, using a ribbon diagram for the enzyme and a ball-and-stick model for the bound amino acid. The polypeptide chain of the enzyme is colour-coded according to the local secondary structure, so that α-helices are coloured red, β-strands blue and random coil green. The active-site zinc cations are shown as pink balls (Zn1 and Zn2, middle left), the bound calcium cation is shown as an orange ball (Ca, top right) and the atoms of the bound inhibitory L-tryptophan (Trp) are coloured indigo. The protein ligands of the zinc ions and the calcium ion are shown in ball-and-stick representation and are coloured according to the usual atom types.

two such amino-acid residues, L-tryptophan (K_i of 5.44 m*M* in the presence of Ca^{2+}; Blumberg *et al.*, unpublished results) and *p*-iodo-L-phenylalanine (K_i of 0.92 m*M* in the presence of Ca^{2+}; Blumberg *et al.*, unpublished results). The detailed binding parameters of the two residues in the active site of the enzyme are presented and the relevance of these interactions to the substrate specificity, catalytic activity and mechanism of action of SGAP is discussed.

2. Materials and methods

2.1. Crystallization

Native SGAP was purified as described previously (Spungin & Blumberg, 1989; Papir *et al.*, 1998). Single crystals of the SGAP–L-Trp complex were grown by the hanging-drop vapour-diffusion method, following the general procedure outlined for the crystallization of the native enzyme (Almog *et al.*, 1993; Greenblatt *et al.*, 1997). The protein solution contained 18 mg ml^{-1} SGAP in 10 m*M* Tris–HCl pH 8.0, 20 m*M* NaCl, 6 m*M* $CaCl_2$ and 100 m*M* L-tryptophan. The protein crystallization drops were made up of equal volumes of the protein solution and a precipitant solution containing 18%(*w*/*v*) PEG 4000 and 0.1 *M* sodium acetate buffer pH 5.5. These drops were equilibrated against a 1 ml reservoir of the precipitant solution for 1 d, followed by seeding with crushed microcrystals of native SGAP. The microseeding procedures were used in order to speed up the crystallization process from about 8–10 weeks to 3–4 d. Single crystals of the SGAP–*p*-iodo-L-phenylalanine complex were grown by a similar procedure using a combination of hanging-drop and vapour-diffusion techniques. The protein crystallization drops were made up of equal volumes of protein solution containing 18 mg ml^{-1} SGAP in 10 m*M* Tris–HCl pH 8.0, 20 m*M* NaCl, 6 m*M* $CaCl_2$ and precipitant solution [18%(*w*/*v*) PEG 4000 and 0.1 *M* acetate buffer pH 5.5] containing 2 m*M* *p*-iodo-L-phenylalanine. A lower concentration is used here because of the higher affinity and lower solubility of *p*-I-L-Phe relative to L-Trp. Microseeding techniques using crushed microcrystals of native SGAP were also used in this case.

2.2. Diffraction data collection

A full high-resolution diffraction data set (to a maximum practical resolution of 1.3 Å) was measured for each of the SGAP complexes (SGAP–L-Trp and SGAP–*p*-I-L-Phe) using a single crystal and synchrotron X-ray radiation. The crystals were flash-frozen in a cold nitrogen-gas stream (95 K) after soaking (for about 2 min) in a cryosolution containing 30%(*w*/*v*) PEG 4000 and 0.1 *M* sodium acetate buffer pH 5.5. Diffraction data from both types of crystals were collected at the National Synchrotron Light Source (NSLS) facility, Brookhaven National Laboratory (BNL) using a Quantum-4 CCD area detector (ADSC, USA). Crystals of the SGAP–L-Trp complex were measured on the X26C/NSLS beamline, while the corresponding crystals of the SGAP–*p*-I-L-Phe complex were measured on the X12B/NSLS beamline. All diffraction data were processed with the *DENZO* and *SCALEPACK* programs (Otwinowski & Minor, 1997) to give the final merged intensity data sets, which were then used for structure determination and refinement. Representative data-collection parameters are listed in Table 1.

Table 1
Summary of crystallographic data-collection and refinement parameters for the complexes of SGAP with L-Trp and *p*-I-L-Phe.

The values for the highest resolution shells are given in parentheses.

	SGAP–L-Trp	SGAP–*p*-I-L-Phe
Data collection		
Space group	$P4_12_12$	$P4_12_12$
Unit-cell parameters		
$a = b$ (Å)	61.34	61.05
c (Å)	144.96	144.50
Resolution range (Å)	20–1.3 (1.32–1.30)	20–1.3 (1.32–1.30)
Collected reflections ($I > 0$)	682173	610801
Unique reflections	62291	63168
Completeness (%)	90.4 (60.2)	92.7 (57.8)
$\langle I/\sigma(I)\rangle$	9.6	17.2
R_{sym}† (%)	9.1 (21.6)	4.5 (16.2)
Refinement		
No. protein atoms	2118	2135
No. inhibitor atoms	30	13
No. water molecules	448	453
R‡ (%)	12.81	12.21
R_{free}‡ (%)	17.42	17.22
R.m.s.d. bonds (Å)	0.012	0.012
R.m.s.d. angles (°)	2.05	2.06

† $R_{sym} = (\sum |I - \langle I\rangle|)/\sum I$, excluding single measurements. ‡ R factor and $R_{free} = (\sum ||F_o| - |F_c||)/\sum |F_o|$ for reflections in the working (R factor) and test (R_{free}) sets, respectively.

2.3. Structure analysis and refinement

The crystals of the two SGAP complexes (with L-Trp and *p*-I-L-Phe) were found to be completely isomorphous to the crystals of native SGAP ($P4_12_12$ space group and less than 2% difference in unit-cell parameters). The structure of native SGAP could therefore be used as the initial model for the structure determination of both complexes. For this purpose, we used the previously reported crystal structure of native SGAP at 1.58 Å resolution (Gilboa *et al.*, 2000; PDB code 1cp7) and the most recent structure, which was refined at 1.35 Å resolution (Gilboa *et al.*, 2004). Crystallographic refinement was initially performed by the *CNS* crystallographic software package (Brünger *et al.*, 1998), subjecting the initial model to several rounds of rigid-body refinement and simulated-annealing minimization. Further improvement of the model was achieved by iterative rounds of model building with the program *O* (Jones *et al.*, 1991), followed by refinement with the program *SHELX*97 (Sheldrick & Schneider, 1997). Solvent molecules were gradually added to the structure from the electron-density difference maps ($F_o - F_c$ coefficients) for difference peaks which were stronger than 3.0σ and located within a proper distance (2.4–3.5 Å) and orientation from a suitable hydrogen-bonding donor or acceptor.

In each of the two cases, the model for the bound amino acid was introduced into the refining model only after all the protein residues had been fully built and most them were well refined. At this point, the difference electron density ($F_o - F_c$)

clearly indicated the location of the bound amino acid and it could easily be placed and oriented in this density (see Fig. 2*a*). The structure of the bound amino acid was then refined together with the protein and the solvent molecules to give a complete structural model. The statistical occupancy of the bound inhibitory amino acid was refined at this stage and converged to a value of 81% for L-tryptophan and a value of 60% for *p*-iodo-L-phenylalanine. The progress of the refinement was monitored *via* $R_{\rm free}$ (Brünger, 1992), which was calculated for a randomly selected set containing 5% of the reflections. After the isotropic refinement of the full model had converged, anisotropic refinement for each of the non-H atoms of the full model took place. The anisotropic refinement significantly improved the geometry of the refined protein and significantly decreased the residual electron density of the maps. The refinement converged to a final *R* factor of 12.81% ($R_{\rm free}$ = 17.42%) for the SGAP–L-Trp complex and a final *R* factor of 12.21% ($R_{\rm free}$ = 17.22%) for the SGAP–*p*-I-L-Phe complex. The overall final structure of the SGAP–L-Trp complex is shown schematically in Fig. 1, in which the secondary-structural elements and the detailed metal environments are demonstrated. Representative refinement parameters for the two structures are given in Table 1.

The matrices for superposition of the different SGAP models were calculated by the least-squares distance-minimization algorithm implemented within the program *O* (Jones *et al.*, 1991) using the C^{α} atoms as the guide coordinates. For the specific comparison of the SGAP active site in the different models, the least-squares calculations included only those residues which are directly involved in zinc coordination (His85, Asp97, Glu132, Asp160 and His247) and enzyme–inhibitor contacts (Glu131, Asp160, Arg202, Phe219 and Tyr246). Figs. 1–3 were prepared using a combination of the programs *MOLSCRIPT* (Kraulis, 1991), *BOBSCRIPT* (Esnouf, 1997) and *RASTER3D* (Merritt & Bacon, 1997).

3. Results and discussion

3.1. The overall structure of the SGAP complexes

The current study used X-ray crystallographic structural analysis of the complexes of SGAP with L-Trp and *p*-I-L-Phe in order to characterize the detailed binding of these product analogues and their interactions with specific residues in the active site of the enzyme. A schematic presentation of the final structure of the SGAP–L-Trp complex is shown in Fig. 1. The overall structure of the corresponding SGAP–*p*-I-L-Phe complex is very similar (not shown). In such cases it is of interest to examine whether conformational changes are induced in the enzyme by the binding of these analogues, especially since they resemble species along the catalytic reaction coordinate. A comparison of the overall structure of the protein involved in these two complexes and the native (uncomplexed) enzyme clearly demonstrates that such binding does not lead to any major structural changes in the enzyme (not shown) and even the local conformational changes around the active site are minimal. Similar results were obtained in the structural analysis of SGAP in its complexes with the smaller amino acids L-Met, L-Leu and L-Phe (Gilboa *et al.*, 2000, 2001).

A least-squares fit of the current structure of the SGAP–L-Trp complex with the most recent structure of native SGAP at 1.35 Å resolution (Gilboa *et al.*, 2004) results in an overall r.m.s. deviation of only 0.25 Å between all of the C^{α} atoms of the two structures. A similar least-squares fit calculation was performed to compare the relevant amino acids of the active site, *i.e.* the residues of the enzyme which are directly involved in zinc coordination and inhibitor binding (see above). This comparison gave an r.m.s. deviation of 0.40 Å for all non-H atoms of these residues and 0.11 Å for the C^{α} atoms of these residues. A similar comparison of native SGAP with the current structure of the SGAP–*p*-I-L-Phe complex results in an overall r.m.s. deviation of 0.28 Å between all of the C^{α} atoms of the two structures and r.ms. deviations of 0.39 and 0.14 Å for all atoms and C^{α} atoms of the active-site residues defined above, respectively. These are definitely very small structural differences and are typical for comparisons of similar structures of relatively rigid proteins that have been determined at comparable resolutions and under similar experimental conditions.

The high overall similarity between the structure of native SGAP and its complexes with relatively large amino-acid residues is very important for the characterization of the active site and the general mode of action of the enzyme. It indicates that the active site of the enzyme is rather open to the solvent, that the binding pocket for the terminal amino acid is relatively large and that no significant conformational changes are involved in the last stages of the catalytic reaction, in which product binding takes place.

In the current structures, several small polypeptide stretches seem to be more flexible than the others, reflected mainly in 'disordered' regions of the protein structure. As in the previously reported structures of SGAP (Greenblatt *et al.*, 1997, Gilboa *et al.*, 2000, 2001), in the present structures there is no interpretable electron density for the last seven amino-acid residues at the C-terminus of the enzyme (residues 278–284). These observations suggest that the carboxy-terminal polypeptide is similarly flexible in both the uncomplexed and the complexed forms of SGAP, indicating that it plays no direct role in either binding or catalysis. In addition to this commonly observed disordered carboxy-terminus, there is another section of the structure of the SGAP–L-Trp complex for which there is no interpretable electron density. This section includes all the atoms (in the main chain and side chain) of residues 198–201 and the side-chain atoms of Arg202. Similarly, in the structure of the SGAP–*p*-I-L-Phe complex there is no interpretable electron density for all the atoms of residues 200 and 201. These regions belong to a flexible loop near the active site (including residues 196–202) which is also shown to be disordered in the structure of the native enzyme at various resolutions and at both room temperature and 95 K (Greenblatt *et al.*, 1997; Gilboa *et al.*, 2000, 2004). The functional significance of this loop and its conformational flexibility still remain to be clarified.

Table 2
Selected enzyme–effector interactions (Å) in the refined crystallographic structures of the complexes of SGAP with L-Trp, *p*-I-L-Phe and L-Phe.

Enzyme atom	Inhibitor atom†	Distance (Å)		
		L-Trp	*p*-I-L-Phe	L-Phe‡
Zn1	OT	2.09	2.20	2.42
Zn2	O	1.84	1.94	2.12
Zn2	OT	2.66	2.55	2.62
Glu131 $O^{\varepsilon 1}$	OT	2.76	2.64	2.80
Tyr246 OH	O	2.82	2.74	2.72
Glu131 $O^{\varepsilon 2}$	N	2.82	2.71	2.98
Asp160 $O^{\delta 2}$	N	2.83	2.74	2.84
Arg202 O	N	2.74	2.78	2.68

† O and OT are the two O atoms of the free carboxylate group. ‡ From Gilboa *et al.* (2001).

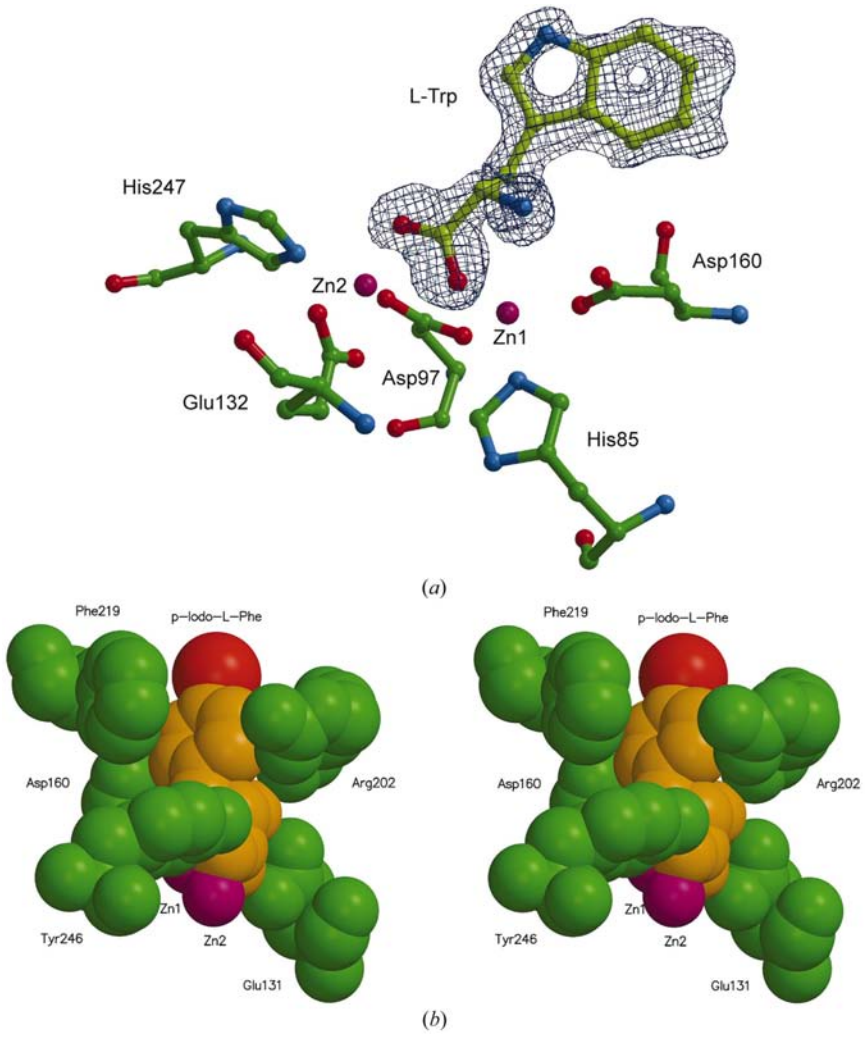

Figure 2
(*a*) Electron-density difference map (omit map, $F_o - F_c$ coefficients) for the final structure of the SGAP–L-Trp complex, demonstrating the clear and unequivocal density (contour level of 1.7σ, cyan) for the bound L-tryptophan (L-Trp) in the active site of SGAP. Also shown are the two active-site zinc ions and their protein ligands. (*b*) Binding of *p*-iodo-L-phenylalanine (*p*-I-L-Phe, orange) in the active site of SGAP (green) as shown in a CPK representation. The active-site Zn atoms are shown as pink balls (bottom, centre) and the I atom of the bound *p*-I-L-Phe is coloured red (top, centre) for clarity. Note the compact contacts between the bound affector and residues of the enzyme in the active site. Also note the aromatic stacking interactions between the phenyl rings of the bound affector and Phe219.

3.2. Enzyme–affector interactions in the active site

3.2.1. L-Tryptophan. The experimental electron density showed very clearly the position and orientation of the bound L-tryptophan in the present crystal structure of the SGAP–L-Trp complex (Fig. 2*a*). The high clarity of the map probably arises from the high resolution of the diffraction data, together with the relatively high crystallographic occupancy (81%) and the limited orientational flexibility of the bound affector. In the refined structure of the SGAP–L-Trp complex, the free tryptophan residue was found to be bound in the active site of the enzyme in a similar fashion to that observed for the other inhibitory amino acids, L-Met, L-Leu and L-Phe (Gilboa *et al.*, 2000, 2001). The most significant interactions of the bound L-Trp with SGAP are made by its free carboxylate group, which contacts the two Zn atoms of the active site and bridges between them (Figs. 2*a* and 3; Table 2). One of these two Zn—O bonds is remarkably short (1.84 Å), indicating a relatively strong Zn—O interaction. The other Zn—O bond (2.09 Å) is more typical for Zn–ligand interactions in biological systems. The tryptophan binding in the active site is also stabilized by a hydrogen bond between one of its carboxylate O atoms and the hydroxyl group of Tyr246 (Fig. 3). On its other side, the tryptophan bridging carboxylate is further stabilized by another hydrogen bond between its second carboxylate O atom and the side-chain carboxylate group of Glu131.

In addition to the interactions of the free carboxylate group of the bound tryptophan residue, there are three hydrogen bonds between its free amine group and neighbouring residues of the enzyme. In these bonds, the amine N atom is the donor and the three acceptors are the main-chain carbonyl O atom of Arg202, the side-chain carboxylate O atom of Asp160 and the other side-chain carboxylate O atom of Glu131. These previously observed hydrogen bonds are unique to SGAP (Gilboa *et al.*, 2001) and seem to be very important for the specific binding of the amino-terminal residue of the target substrates (and the cleaved product). These interactions contribute to the specific function of SGAP as an aminopeptidase. The preference of the enzyme for N-terminal amino acids that are hydrophobic and/or aromatic is generally accounted for by the relatively large specificity pocket that is surrounded by hydrophobic/aromatic side chains of the enzyme. In the present case, the indole ring of the tryptophan side chain is well fitted within this hydrophobic pocket. Also noticed here

is an enthalpically favourable 'face-to-face' aromatic interaction (Burley & Petsko, 1985, 1986) between the six-membered ring of the tryptophan side chain and the phenyl ring of Phe219 of the enzyme. The angle between the two general planes of these aromatic rings is about 14° and the distance between the centres of the rings is about 3.9 Å. A similar aromatic interaction is observed in the structure of the SGAP–*p*-I-L-Phe complex (Fig. 2*b*), as discussed below.

In addition to the L-tryptophan molecule that was found tightly bound in the active site of SGAP, a second L-tryptophan molecule was found on the surface of the enzyme (with a crystallographic occupancy of about 60%) at a distance of about 22 Å from the active site. This L-tryptophan is bound to the outer layer of the protein, mainly through its free amine and carboxylate groups. The amine N atom donates two hydrogen bonds, one to the carboxylate O atom of Asp258 and the other to the carbonyl O atom of Ile256. Similarly, the two carboxylate O atoms of the bound tryptophan accept hydrogen bonds from the amine N atom of Asp258. This is the first time that a surface-binding site has been observed for SGAP (or for any other aminopeptidase) and it is as yet unclear whether it has any functional significance. Alternatively, such binding could be a consequence of the relatively high concentration of L-Trp (100 m*M*) used for the cocrystallization experiments.

3.2.2. *p*-Iodo-L-phenylalanine. In the structure of the SGAP–*p*-I-L-Phe complex, the position and the interactions of the bound amino acid in the active site of the enzyme (with an occupancy of about 60%) are very similar to those of L-tryptophan in the SGAP–L-Trp complex described above. The close contacts of the bound effector with residues in the active site of the enzyme are demonstrated in Fig. 2(*b*), where the relevant moieties are shown in a CPK representation. Here, the free carboxylate group of the bound amino acid again acts as a bidentate ligand to one of the two Zn atoms, while only one of the two carboxylate O atoms bridges between the two Zn atoms. Further stabilization is achieved on the two sides of the zinc-bound carboxylate by a net of hydrogen bonds, with the hydroxyl group of Tyr246 on one side and with the carboxylate group of Glu131 on the other. The geometry of this network is very similar to that of the SGAP–L-Trp complex, as described above (Fig. 3; Table 2) and seems to be a common and important feature in binding and catalysis employed by SGAP and related enzymes (Gilboa *et al.*, 2001).

The only obvious difference in the orientation of the L-Trp and the p-I-L-Phe in their complexes with SGAP is a slight conformational change in the *p*-iodo-L-phenylalanine side chain (a rotation of about 32° around the $C^{\alpha}-C^{\beta}$ bond) relative to the corresponding conformation of the side chain in the L-tryptophan complex. This conformational change forces the phenyl aromatic ring of the bound *p*-iodo-L-phenylalanine into a slightly less favourable interaction with Phe219 of the enzyme, with an angle of about 35° between the two ring planes (the distance between the centres of the rings is about 4.5 Å). This conformational change and the slight displacement of the bound amino acid is probably a result of the abnormal size and length of the side chain owing to its *p*-iodo substituent.

It should be noted here that a weak, yet significant, peak of residual electron density was still present in the final rounds of refinement of the SGAP–*p*-I-L-Phe complex. This extra electron density ($F_o - F_c$) is located very close to the current location of the bound amino acid and may indicate an alternative (and significantly weaker) binding geometry of this modified amino acid in the active site of SGAP.

3.3. Binding of bulky aromatic amino acids to SGAP

It is interesting to compare the exact binding geometry and the detailed enzyme–analogue interactions for the SGAP complexes with the closely related inhibitory amino acids L-Trp, *p*-iodo-L-Phe (this work) and L-Phe (Gilboa *et al.*, 2001). It is expected that common features will indicate the general substrate binding and the general catalytic mode of action of SGAP, while differences and non-common features are likely to point out aspects related to substrate (or product) specificity and specificity-related mechanistic factors, if any.

As already noted above, the overall structure of the enzyme remains practically unchanged, even after the binding of the bulky amino acids L-Trp and *p*-iodo-L-Phe, and the same is true for L-Phe. A superposition of the C^{α} coordinates of the three complexes (not shown here) indicates no significant differences when comparing the overall enzyme structure. Similarly, the overall conformation of SGAP in the three complexes shows no significant changes compared with the native enzyme (1.35 Å resolution; Gilboa *et al.*, 2004). The r.m.s.d. in the backbone coordinates between the native

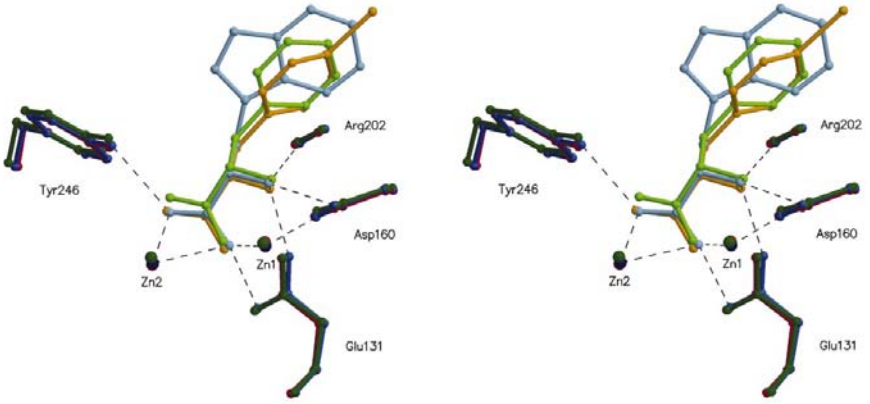

Figure 3
Stereoview comparison of the active site of the SGAP–L-Trp complex (1.3 Å resolution, blue) with the active site of the SGAP–*p*-I-L-Phe complex (1.3 Å resolution, orange) and the active site of the SGAP–L-Phe complex (1.8 Å resolution, green; Gilboa *et al.*, 2001). The corresponding bound affectors are shown in light blue, light orange and light green, respectively. Note the very similar binding geometry of the carboxylate and the amine groups of the three bound affectors to the active-site zinc ions and neighbouring residues of the enzyme.

enzyme and the three enzyme complexes is less than 0.3 Å, demonstrating quantitatively that the structure of the protein backbone is not affected by analogues of this size and chemical nature. The small differences in the unit-cell parameters of these three SGAP–amino acid complexes (Table 1; $a = b = 62.18$, $c = 146.98$ Å for the L-Phe complex) are likely to be a consequence of subtle structural changes in each particular complex. These changes could result from the slightly different solvent content and enzyme–solvent interactions, as well as slightly different experimental factors, such as the exact crystallization conditions and the exact crystallographic data-measurement procedures. These results are not obvious and demonstrate that the small differences observed in the orientation of the bound analogues (see below), as well as the small changes in their binding and inhibition constants, arise from specific interactions in the active site of the enzyme and not from indirect effects that lead to global conformational changes.

A detailed comparison of the active sites of the SGAP–L-Trp, SGAP–*p*-I-L-Phe and SGAP–L-Phe complexes shows that the binding modes of all three amino acids are very similar, despite the differences in size, shape and polarity of the three side chains (Fig. 3). The position and orientation of the L-phenylalanine and the *p*-iodo-L-phenylalanine appear to be almost identical (Fig. 3), except for a slight conformational change in the side chains (a rotation of about 11° around the $C^{\beta}-C^{\gamma}$ bond) and a slight shift of the *p*-iodo-L-phenylalanine molecule towards the two Zn atoms, resulting in relatively short Zn—O interactions (Table 2). A comparable shift is also observed in the active site of the SGAP–L-Trp complex. Another difference in the position and orientation of the bound affectors discussed above is a rotation of about 28° of the side chain of L-Trp around its $C^{\alpha}-C^{\beta}$ bond in comparison to the corresponding orientation of the side chain of the bound L-phenylalanine.

For all three bound inhibitory amino acids compared here, there are stacking interactions between the aromatic ring of Phe219 and the aromatic rings of the bound analogues (as demonstrated in Fig. 2*b* for *p*-I-L-Phe). The bound L-phenylalanine and the L-tryptophan show only slight deviations from the enthalpically favourable 'face-to-face' interaction (Burley & Petsko, 1985, 1986). The angles between the corresponding aromatic rings are 19 and 14° and the distances are 4.3 and 3.9 Å, respectively. In all three cases, however, there are no significant changes in the position or orientation of the aromatic side chain of the enzyme residue Phe219.

Except for the residues in the flexible loop of the enzyme (196–202), for which there was no interpretable electron density (and their conformation is therefore undefined), no significant conformational changes of active-site residues were observed in the structures of the three SGAP complexes. This observation again demonstrates that the enzyme can accommodate a relatively wide range of amino acids and their analogues without the need for conformational changes or related structural adaptation. The absence of significant conformational changes and the relatively open binding space of the active site may partially account for the relatively poor inhibition values (millimolar range) obtained for all the inhibitory amino acids examined, despite their tight interactions with the double-zinc centre.

The small differences in the exact orientation of the three aromatic amino acids within the active site of SGAP is somewhat surprising considering the significant differences in their corresponding K_i values (12.7, 5.44 and 0.92 m*M* for L-Phe, L-Trp and *p*-I-L-Phe, respectively). On the basis of the structural data presented above, it seems that the order of the K_i values is influenced by a combination of several binding interactions. The exact geometry of one of these interactions, the aromatic 'face-to-face' stacking, appears to be less critical for the overall binding. The increased hydrophobic/aromatic character and the overall size of the side chain alone is probably one of the more significant binding factors, as reflected by the L-Phe *versus* L-Trp comparison. The presence and exact location of the bulky and polarizable I atom at the *para* position of the phenyl ring seems to be even more significant for binding interactions in the active site of SGAP, as demonstrated by the L-Phe *versus* *p*-I-L-Phe comparison. The quantitative contribution of each of these specific binding factors should be further examined by detailed molecular simulations.

3.4. Mechanistic interpretations

In order to use the present structural data to gain insight into the catalytic mode of action of SGAP, it is most important to compare the detailed interactions of the three inhibitory amino acids with the active-site zinc ions and the mechanistically relevant residues around them. In this respect, the three residues show very similar binding geometry. All three were found to be bound to the active-site zinc ions, displacing the zinc-bound water/hydroxide of the native enzyme in the solvent-accessible face of the active site. The free carboxylate group that bridges the two zinc ions makes the most significant contacts of the bound amino acids with SGAP. These two Zn—O interactions are relatively short, yet they are not symmetrical (Table 2). As observed previously for L-Met and L-Leu, the binding of each of the amino acids examined here is further stabilized by a hydrogen bond between one of its carboxylate O atoms and the hydroxyl group of Tyr246, and a hydrogen bond between the other carboxylate oxygen and the carboxylate group of Glu131. In addition, for all the amino-acid complexes examined so far, there are three hydrogen bonds between the amino-acid amine N atom (probably present as an $-NH_3^+$ group) and the backbone carbonyl O atom of Arg202, a carboxylate O atom of Asp160 and the other carboxylate O atom of Glu131.

The high degree of similarity of these interactions, regardless of the type of the bound amino acid and its specific orientation in the active site, suggests that they are important for the proper mode of action of the enzyme. In this respect, it should be noted that all the amino acids examined above act as weak inhibitors of SGAP, with inhibition constants in the millimolar range. They can potentially be regarded also as product analogues, since one of the products obtained in a

normal enzymatic reaction is the cleaved amino-terminal amino acid of the substrate. Moreover, since all the amino acids bind to the zinc ions with their free carboxylate group in a bidentate fashion, they can also serve, at least in part, as models of earlier species along the reaction coordinate. Such species may, for example, be the first intermediate (or transition state) of the catalytic reaction. This species results from the initial nucleophilic attack of the bound water/hydroxide on the carbonyl group of the substrate and is assumed to take the form of a tetrahedral *gem*-diolate moiety (Gilboa *et al.*, 2001). The structural features discussed above could therefore also be relevant for these earlier catalytic species, emphasizing the role of the zinc ions in binding, orienting and neutralizing such charged species, as well as generalizing the role of Tyr246 and Glu131 in the stabilization of these species. The involvement of these elements in the detailed catalytic mechanism of SGAP has already been suggested (Gilboa *et al.*, 2001) and is fully supported here. Similar *gem*-diolate species have been proposed as the tetrahedral catalytic intermediate of the mono-zinc proteinase carboxypeptidase A (CPA; Shoham *et al.*, 1988; Feinberg *et al.*, 1993, 1995) which also contains a catalytically critical zinc-bound water/hydroxide (Shoham *et al.*, 1984). For CPA it was also proposed that this intermediate is stabilized by the carboxylate group of a nearby glutamic acid residue (Glu270) (Feinberg *et al.*, 1995; Kilshtain-Vardi *et al.*, 2002) similar to Glu131 in SGAP. Both glutamates are suggested to act as general acids in the initial catalytic steps.

4. Conclusions

The high-resolution crystal structures described in the present work show that the binding to SGAP of the bulky amino acids L-tryptophan and *p*-iodo-L-phenylalanine is very similar and that their general mode of binding is similar also to the previously studied amino acids Met, Leu and Phe. Several conclusions can be drawn from these results, some of which are directly applicable to the catalytic mechanism of SGAP.

(i) Binding of amino acids to the active site of SGAP does not alter the structure of the unbound enzyme.

(ii) The specificity pocket of SGAP is limited in size and shape, but allows the binding of a relatively wide range of side chains within these limits.

(iii) Amino acids use their free carboxylate group to bind directly to the two active-site zinc ions, while displacing the zinc-bound water of the native enzyme.

(iv) The free carboxylate of the bound residue is further stabilized by hydrogen bonds to Tyr246 and Glu131, both of which seem to play a critical role in the catalytic reaction.

(v) The free amine group of the bound amino acid is involved in specific hydrogen bonds with three specific protein residues in the active site, forming a dedicated 'amino-terminus binding subsite', as previously suggested.

(vi) In the case of aromatic amino acids such as Phe and Trp, the binding of the aromatic side chain within the specificity pocket involves also an important stacking interaction with the phenyl ring of Phe219, accounting for the high preference of SGAP for N-terminal aromatic residues.

These and related conclusions from the present study support the general mechanism for the catalytic reaction proposed earlier for SGAP (Gilboa *et al.*, 2001). This support is important since the corresponding mechanisms proposed for related aminopeptidases (such as AAP and LAP) appear to be significantly different in both the binding geometry and the course of the catalytic reaction. Obviously, these differences should be validated and clarified by further systematic studies, especially those that will combine structural analysis, site-directed mutagenesis and kinetics.

We would like to thank the National Synchrotron Light Source (NSLS), the biology department of Brookhaven National Laboratory, and Dr Dieter Schneider for beam-time allocation and assistance at the X12B and X26C beamlines. This work was supported in part by a grant from the German Federal Ministry of Education, Science, Research and Technology (BMBF) and the Israeli Ministry of Science (MOS) under the aegis of the GBF – Gesellschaft fur Biotechnologische Forschung GmbH, Braunschweig. It was also supported by the Israel Academy of Sciences and Humanities–Israel Science Foundation (ISF). VR was supported by a Marie Curie Fellowship of the European Commission and by a Lady Davis Fellowship from the Hebrew University of Jerusalem, Israel.

References

Almog, O., Greenblatt, H. M., Spungin, A., Ben-Meir, D., Blumberg, S. & Shoham, G. (1993). *J. Mol. Biol.* **230**, 342–344.
Ben-Bassat, A., Bauer, K., Chang, S. Y., Myambo, K., Boosman, A. & Chang, S. (1987). *J. Bacteriol.* **169**, 751–757.
Ben-Meir, D. & Blumberg, S. (1991). *Biologicals from Recombinant Microorganisms and Animal Cells: Production and Recovery*, edited by M. D. White, S. Reuveny & A. Shaffermann, pp. 55–61. Weinheim, Germany: Verlag Chemie.
Ben-Meir, D., Spungin, A., Ashkenazi, R. & Blumberg, S. (1993). *Eur. J. Biochem.* **212**, 107–112.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
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., David, P. R., Sweet, R. M., Taylor, A. & Lipscomb, W. N. (1992). *J. Mol. Biol.* **224**, 113–140.
Burley, S. K. & Petsko, G. A. (1985). *Science*, **229**, 23–28.
Burley, S. K. & Petsko, G.A. (1986). *J. Am. Chem. Soc.* **108**, 7995–8001.
Chevrier, B., Schalk C., D'Orchymont, H., Rondeau, J.-M., Moras, D. & Tarnus, C. (1994). *Structure*, **2**, 283–291.
Esnouf, R. M. (1997) *J. Mol. Graph.* **15**, 132–138.
Feinberg, H., Greenblatt, H. M., Rashti, V., Gilon, C., Cohen, S., Bino, A. & Shoham, G. (1995). *Acta Cryst.* D**51**, 428–449.
Feinberg, H., Greenblatt, H. M. & Shoham, G. (1993). *J. Chem. Inf. Comput. Sci.* **33**, 501–516.
Fundoiano-Hershcovitz, Y., Rabinovitch, L., Langut, Y., Reiland, V., Shoham, G. & Shoham, Y. (2004). *FEBS Lett.* **571**, 192–196.
Gilboa, R., Greenblatt, H. M., Perach, M., Spungin-Bialik, A., Lessel, U., Wohlfahrt, G., Schomburg, D., Blumberg, S. & Shoham, G. (2000). *Acta Cryst.* D**56**, 551–558.
Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S. & Shoham, G. (2001). *Proteins*, **44**, 490–504.

Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S. & Shoham G. (2004). In preparation.
Greenblatt, H. M., Almog, O., Maras, B., Spungin-Bialik, A., Barra, D., Blumberg, S. & Shoham, G. (1997). *J. Mol. Biol.* **265**, 620–636.
Indig, F. E., Ben-Meir, D., Spungin, A. & Blumberg, S. (1989). *FEBS Lett.* **255**, 237–240.
Indig, F. E., Benayahu, D., Fried, A., Wientroub, S. & Blumberg, S. (1990). *Biochem. Biophys. Res. Commun.* **172**, 620–626.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kilshtain-Vardi, A., Shoham, G. & Goldblum, A. (2002). *Int. J. Quant. Chem.* **88**, 87–98.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Lawrence, C. M., Ray, S., Babyonyshev, M., Galluser, R., Borhani, D. W. & Harrison, S. C. (1999). *Science*, **286**, 779–782.
Mahadevan, D. & Saldanha, J. W. (1999). *Protein Sci.* **8**, 2546–2549.
Maras, B., Greenblatt, H. M., Shoham, G., Spungin, A., Blumberg, S. & Barra, D. (1996). *Eur. J. Biochem.* **236**, 843–846.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Nagakawa, S., Yamada, T., Kato, K. & Nishimura, O. (1987). *Biotechnology*, **5**, 824–827.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Papir, G., Spungin-Bialik, A., Ben-Meir, D., Fudim, E., Gilboa, R., Greenblatt, H. M., Shoham, G., Lessel, U., Schomburg, D., Ashkenazi, R. & Blumberg, S. (1998). *Eur. J. Biochem.* **258**, 313–319.
Roderick, S. L. & Matthews, B. W. (1993). *Biochemistry*, **32**, 3907–3912.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Shoham, G., Christianson, D. W. & Oren, D. A. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 684–688.
Shoham, G., Rees, D. C. & Lipscomb, W. N. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 7767–7771.
Spungin, A. & Blumberg, S. (1989). *Eur. J. Biochem.* **183**, 471–477.
Taylor, A. (1993*a*). *FASEB J.* **7**, 290–298.
Taylor, A. (1993*b*). *Trends Biochem. Sci.* **18**, 167–172.
Taylor, A. (1996). Editor. *The Aminopeptidases*, pp. 1–20. Austin, TX, USA: Landes Bioscience Publishers.
Wilce, M. C., Bond, C. S., Dixon, N. E., Freeman, H. C., Guss, J. M., Lilley, P. E. & Wilce, J. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 3472–3477.