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

Structure of *Bacillus amyloliquefaciens* α-amylase at high resolution: implications for thermal stability

Jahan Alikhajeh,[a,b] Khosro Khajeh,[b]* Bijan Ranjbar,[c] Hossein Naderi-Manesh,[c] Yi-Hung Lin,[a] Enhung Liu,[a] Hong-Hsiang Guan,[a] Yin-Cheng Hsieh,[a] Phimonphan Chuankhayan,[a] Yen-Chieh Huang,[a] Jeyakanthan Jeyaraman,[a] Ming-Yih Liu[a,d] and Chun-Jung Chen[a,e,f]*

[a]Life Science Group, Scientific Research Division, National Synchrotron Radiation Research Center, Hsinchu 30076, Taiwan, [b]Department of Biochemistry, Faculty of Science, Tarbiat Modares University, Tehran, Iran, [c]Department of Biophysics, Faculty of Science, Tarbiat Modares University, Tehran, Iran, [d]Institute of Nuclear Energy Research Center, Atomic Energy Council, Taoyuan 32546, Taiwan, [e]Department of Physics, National Tsing Hua University, Hsinchu 30014, Taiwan, and [f]Institute of Biotechnology, National Cheng Kung University, Tainan 70101, Taiwan

Correspondence e-mail: khajeh@modares.ac.ir, cjchen@nsrrc.org.tw

Received 14 June 2009
Accepted 2 December 2009

PDB Reference: *Bacillus amyloliquefaciens* α-amylase, 3bh4.

The crystal structure of *Bacillus amyloliquefaciens* α-amylase (BAA) at 1.4 Å resolution revealed ambiguities in the thermal adaptation of homologous proteins in this family. The final model of BAA is composed of two molecules in a back-to-back orientation, which is likely to be a consequence of crystal packing. Despite a high degree of identity, comparison of the structure of BAA with those of other liquefying-type α-amylases indicated moderate discrepancies at the secondary-structural level. Moreover, a domain-displacement survey using anisotropic *B*-factor and domain-motion analyses implied a significant contribution of domain B to the total flexibility of BAA, while visual inspection of the structure superimposed with that of *B. licheniformis* α-amylase (BLA) indicated higher flexibility of the latter in the central domain A. Therefore, it is suggested that domain B may play an important role in liquefying α-amylases, as its rigidity offers a substantial improvement in thermostability in BLA compared with BAA.

1. Introduction

α-Amylase [EC 3.2.1.1; α-(1,4)-D-glucan glucanohydrolase] is an *endo*-type enzyme that hydrolyzes starch and glycogen by cleaving α-1,4-glucosidic linkages in a random fashion. α-Amylases are widely distributed in various bacteria, fungi, plants and animals and play major roles in the utilization of polysaccharides. Among them, those from *Bacillus* have been intensively studied and are important in industry. The nucleotide sequences of α-amylases from *B. amyloliquefaciens* (BAA), *B. stearothermophilus* (BstA), *B. licheniformis* (BLA) and other *Bacillus* species have been determined (Takkinen *et al.*, 1983; Ihara *et al.*, 1985; Nakajima *et al.*, 1985; Yuuki *et al.*, 1985; Yang *et al.*, 1983; Yamazaki *et al.*, 1983). BAA, BStA and BLA are liquefying-type enzymes and their primary structures resemble each other. The amino-acid sequence of BLA shows 80% identity to that of BAA and 65% to that of BStA (Yuuki *et al.*, 1985). Despite the resemblance in their primary structures, these three enzymes exhibit diverse stabilities towards heat and acidity: their thermal stability increases greatly in the order BAA, BStA, BLA. BLA shows a half-life that is more than 100 times greater than that of BAA at 363 K (Tomazic & Klibanov, 1988). Various mutational approaches have been undertaken in order to determine the amino-acid residues that are responsible for the specificity and thermostability of *Bacillus* α-amylases (Smirnova *et al.*, 1987; Suzuki *et al.*, 1989; Holm *et al.*, 1990; Vihinen *et al.*, 1990; Svensson, 1991; Svensson & Sogaard, 1992; Nagashima *et al.*, 1992; Takase *et al.*, 1992). As attempts to find a suitable crystal for structural determination of BAA have failed (Suzuki *et al.*, 1990; Walker & Campbell, 1967*a*,*b*), most of the effort towards explaining the thermostability at the structural level has been restricted to the crystal structures of hyperthermostable homologues (BLA and BStA). Here, we have succeeded in determining the crystal structure of BAA at high resolution, which we compared with those of hyperthermostable homologues, scrutinizing the mutational and experimental data at the structural level. Moreover, we have analyzed four regions that have previously been proposed to be responsible for thermostability.

2. Experimental procedures

2.1. Purification and crystallization

α-Amylase (20 ml) from *B. amyloliquefaciens* (Sigma Co.) was dialyzed overnight against excess 20 m*M* Tris–HCl pH 7.5. A crude extract with little ionic content was loaded onto a DEAE-Sepharose column (120 ml gel) and a linear gradient of NaCl (0–400 m*M*) was applied. All fractions containing α-amylase were concentrated after screening and were pooled. The active α-amylase fractions were loaded onto a column (G-50, dimensions 150 × 2 cm) for further purification. 20 m*M* Tris–HCl pH 7.5 containing 150 m*M* NaCl was applied onto the pre-equilibrated column at a flow rate of 1 ml min^{-1}. The purity of the active fractions was tested with SDS–PAGE. Prior to crystallization, BAA was extensively dialyzed against 10 m*M* Tris–HCl pH 7.5 containing 1 m*M* $CaCl_2$ and concentrated to 15 mg ml^{-1}. Crystallization was performed in VDX24 plates (Hampton Research) using the hanging-drop vapour-diffusion method at 291 K. Small crystals were observed in a condition consisting of 30%(*w*/*v*) PEG 3350 and 3.5 m*M* calcium chloride in 10 m*M* Tris buffer pH 7.5 within 15 d of setup. This condition was improved to produce larger BAA crystals using hanging drops (2 µl) containing equal volumes of protein solution and reservoir solution equilibrated against 200 µl reservoir solution containing 24%(*w*/*v*) PEG 3350 and 3.5 m*M* $CaCl_2$ in 10 m*M* Tris buffer pH 7.5.

2.2. X-ray data collection and processing

The protein crystals were initially tested and characterized using synchrotron radiation on SPXF beamlines BL13B1 and BL13C1 equipped with CCD detectors (Q315 and Q210, ADSC) at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. Crystals of satisfactory diffraction quality were used for data collection. A crystal was transferred from a crystallization drop into a cryoprotectant solution (5 µl) containing 24%(*w*/*v*) PEG 3350, 3.5 m*M* $CaCl_2$ and 20%(*v*/*v*) glycerol in 10 m*M* Tris buffer pH 7.5 for a few seconds, mounted on a synthetic nylon loop (0.1–0.2 mm, Hampton Research) and flash-cooled in liquid nitrogen at a temperature of 70 K. For complete data collection, 360° of rotation was measured with 0.5° oscillations using an X-ray wavelength of 1.00 Å, an exposure duration of 30 s and a crystal-to-detector distance of 200 mm at 100 K in a dinitrogen stream using a cryosystem (X-Stream, Rigaku/MSC Inc.). All data were indexed, integrated and scaled using programs from the *HKL*-2000 suite (Otwinowski & Minor, 1997). Analysis of the diffraction pattern indicated that the crystals belonged to space group $P2_12_12$, with unit-cell parameters a = 89.76, b = 148.68, c = 76.20 Å. The data set was 98.3% complete with an internal agreement (R_{merge}) of 6.5%. Assuming the presence of two molecules per asymmetric unit, the Matthews coefficient was estimated to be 2.5 $Å^3$ Da^{-1}, which corresponds to a solvent content of 51.6% (Matthews, 1968). Details of the data statistics are given in Table 1. All graphical work and structure comparison were performed using the *PyMOL* program (DeLano, 2002).

2.3. Crystal structure determination, refinement and domain-motion analysis

The crystal structure of BAA was solved by molecular replacement (Rossmann, 1990) with *CNS* v.1.1 (Brünger *et al.*, 1998), using the chimeric BAA/BLA monomer structure (PDB entry 1e43; Brzozowski *et al.*, 2000) as a search model. The rotation- and translation- function searches were calculated using data in the resolution range 30–3 Å and a Patterson radius of 25 Å, which gave a unique solution with a large correlation between the observed amplitudes for the crystal and the calculated amplitudes for the model. Further crystallographic refinement was performed with *CNS* v.1.1. A random selection (5%) of the data was set aside as a 'free data set' throughout refinement and the model was refined against the rest of data as the working data set (Brünger, 1992). The protein model obtained by molecular replacement was initially refined using data from 30.0 to 1.4 Å resolution. The individual B values were first restrained to 20.0 $Å^2$ and were only refined in the last cycles. After rigid-body refinement the solution attained an initial R factor of 0.315 and an R_{free} of 0.314 in the resolution range 30–1.4 Å. This refinement was followed by simulated annealing with a slow-cooling protocol in *CNS* applied to data between 30.0 and 1.4 Å resolution. The temperature was gradually decreased from 2500 to 300 K in steps of 25 K using a time step of 0.5 fs between energy calculations. The R factor decreased to 0.274 and R_{free} decreased to 0.281 for this resolution range. Composite maps with $|2F_o - F_c|$ coefficients were calculated with *CNS* and visualized using the *NOC* program (Chen *et al.*, 2007) throughout refinement and the model was built and adjusted iteratively as required. In the later stages of refinement, a bulk-solvent correction was applied; individual B factors were adjusted, giving an R factor of 0.270 and an R_{free} of 0.279. The *PICKWATER* subroutine from *CNS* served to define maxima in $|F_o - F_c|$ difference maps (3σ cutoff level) for the location of water molecules; a water molecule was accepted if the identified maximum correlated with a separate maximum in the corresponding $|2F_o - F_c|$ electron-density map

Table 1
Crystal parameters, data-collection and structure-refinement statistics.

Values in parentheses are for the highest resolution shell.

Data statistics	
Wavelength (Å)	1.0
Space group	$P2_12_12$
Unit-cell parameters	
a (Å)	89.76
b (Å)	148.68
c (Å)	76.20
Resolution limits (Å)	30.0–1.4 (1.45–1.40)
Unique reflections	197183 (19589)
Completeness (%)	99.5 (100)
Mean redundancy	7.1 (7.1)
R_{merge}† (%)	6.5 (46.3)
$\langle I/\sigma(I)\rangle$	27.18 (4.52)
Model statistics	
No. of reflections in refinement	196849
No. of reflections for R_{free} calculation	9917
R factor‡	16.3
R_{free}§ (%)	20.7
No. of protein non-H atoms	7778
No. of calcium ions	8
No. of sodium ions	2
No. of water molecules	647
Wilson B factor ($Å^2$)	12.7
Average B factor ($Å^2$)	15.52
Protein atoms	14.70
Calcium ions	11.41
Sodium ions	11.76
Water molecules	23.47
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	1.32
Dihedral angles (°)	23.25
Ramachandran plot	
Most favoured region (%)	86.8
Additional allowed regions (%)	12.4
Generously allowed regions (%)	0.5
Disallowed regions (%)	0.2

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith measurement and $\langle I(hkl)\rangle$ is the weighted mean of all measurements of $I(hkl)$. ‡ R factor = $\sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes of reflection hkl, respectively. § The R_{free} factor was calculated using 5% of the unique reflections.

Table 2
Comparison of different properties of liquefying-type α-amylases from different sources.

PDB code	R.m.s.d. for main chain (Å)	R.m.s.d. for all atoms (Å)	Identity for aligned residues (%)	Similarity (%)	Overall accessible surface area (Å^2)	Total electrostatic potential charge
3bh4 (BAA)	0	0	100	100	17642.95	−16
1e43 (BA2)	0.37	0.58	93.37	95.77	17674.10	−15
1bli (BLA)	0.65	0.74	80.24	87.35	17753.82	−12
1hvx (BStA)	0.95	0.95	65.41	73.06	17724.63	−11

and if one or more hydrogen bonds (3.3–2.3 Å) were identifiable. According to these criteria, 647 water molecules were located. The protein model and water molecules were then subjected to another run of positional, simulated-annealing and individual *B*-factor refinement. For data between 30.0 and 1.4 Å resolution, the resulting model had a final *R* factor of 20.6% and an $R_{\rm free}$ of 21.9% (Table 1).

Further refinement was carried out using *CCP*4 (Collaborative Computational Project, Number 4, 1994) and programs supported therein. Firstly, restrained refinement was performed in *REFMAC* (Murshudov *et al.*, 1997) for ten cycles using conjugate-gradient minimization against a least-squares target function with stereochemical restraints, leading to a reduction in *R* and $R_{\rm free}$ from 0.206 and 0.219 to 0.163 and 0.207, respectively. Prior to refinement, all atomic temperature factors were set to a constant value of 20 Å^2. A bulk-solvent correction model with the parameters of the mask optimized to 1.8 Å for VDWProb, 1.8 Å for IONProb and 1.1 Å for RSHRink and overall anisotropic scaling was applied. This model was then used for refinement in unrestrained mode and anisotropic temperature factors using *CCP*4, giving an *R* factor of 16.3% and and an $R_{\rm free}$ of 20.7%. Riding H atoms (C—H and N—H) were located automatically by *CCP*4 and their coordinates were not refined. Finally, the *TLSMD* server (http://skuld.bmsc.washington.edu/~tlsmd) was used to interpret the anisotropic temperature factor in terms of several groups (Painter & Merritt, 2006). For each protein chain, this analysis showed the optimal division into one group, two groups, three groups *etc.* up to 20 groups. It continued to analyze the implied rigid-body translational and rotational motion of each group as well as its quality of fit to the refined atomic displacement parameters (*B* factors).

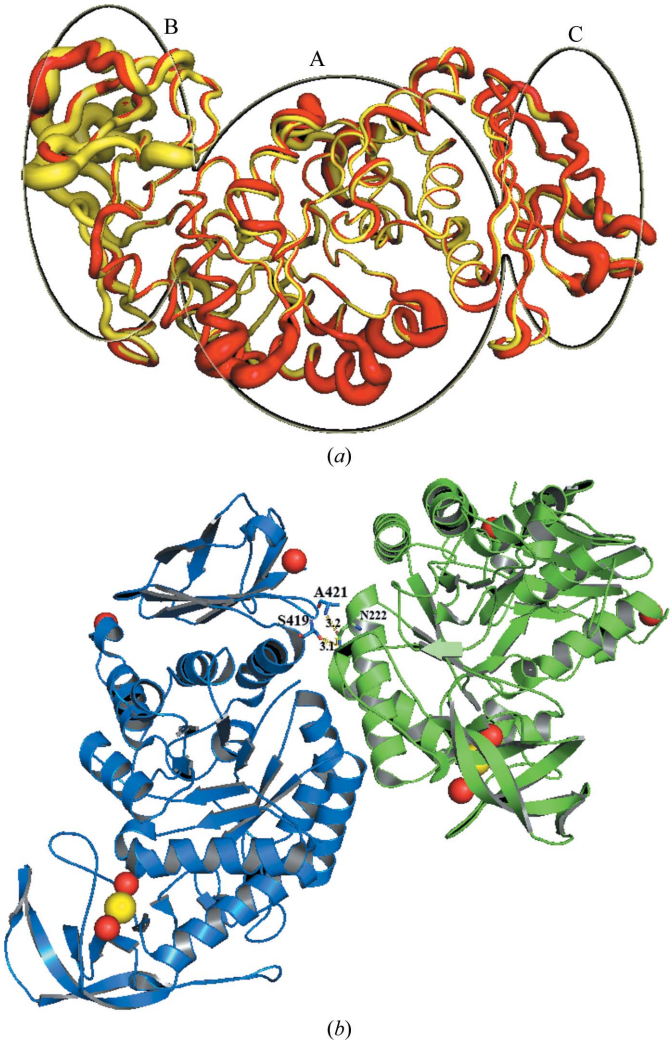

Figure 1
Superimposition of BAA (yellow) and BLA (red) structures; the thickness of the coils shows their corresponding isotropic *B* factors. The ellipses indicate the three distinct domains in amylases named A, B and C. (*a*) The final model composed of two monomers (chain *A*, left; chain *B*, right) in a back-to-back orientation with respect to each other. (*b*) Interaction between chains *A* and *B* in the dimeric form through hydrogen bonding of Asn222 from chain *A* to the nitrogen and carbonyl group of chain *B*; calcium and sodium ions are presented as red and yellow spheres, respectively. Distances are specified in Å.

The correctness of the stereochemistry of the model was verified using *PROCHECK* (Laskowski *et al.*, 1993) and r.m.s. deviations from ideal values for bonds, angles, dihedral angles and improper angles were calculated in *CNS*. For all criteria used by *PROCHECK*, the model was flagged as being better than or within satisfactory regions. The calculations showed also satisfactory stereochemistry, with r.m.s. deviations of 0.005 Å from ideal bond lengths (Engh & Huber, 1991) and 1.32° from ideal bond angles. In a Ramachandran plot (Ramachandran & Sasisekharan, 1968), 98.2% of residue dihedral angles occurred in the most favoured and additionally allowed regions. The occurrence of Tyr149 in the disallowed region is characteristic of α-amylases and has also been commented on for the BL2 and BLA structures (Brzozowski *et al.*, 2000; Machius *et al.*, 1998).

2.4. Structural alignment and comparison

Protein coordinates of liquefying-type α-amylases were obtained from the RCSB Protein Data Bank. Primary sequence alignment was performed with *ClustalW* (Thompson *et al.*, 1994). The *NOC* software was employed to obtain the accessible surface area, for r.m.s.d. calculations and to determined the percentage identity/similarity (Chen *et al.*, 2007). The secondary-structure assignments and other graphical representations in Fig. 2 were produced using the *ESPript* program (Gouet *et al.*, 1999).

3. Results and discussion

3.1. Overall topology of BAA and comparison of sequence alignment

Since a detailed structural study of all known α-amylases is beyond the scope of this contribution, we chose four crystal structures of closely related liquefying-type α-amylases in order to conduct a comparison of their differences in thermostability. The amino-acid sequences of these structures range from 483 to 515 residues in

length. The structure of BAA contains three distinct domains (A, B and C) that have outstanding similarity to those of other thermophilic liquefying-type bacterial α-amylases (Fig. 1*a*). Judging from r.m.s.d. and sequence identity (Table 2), the structure of BAA is highly similar to that of the previously solved chimeric structure (Brzozowski *et al.*, 2000). Domain A is a central well conserved

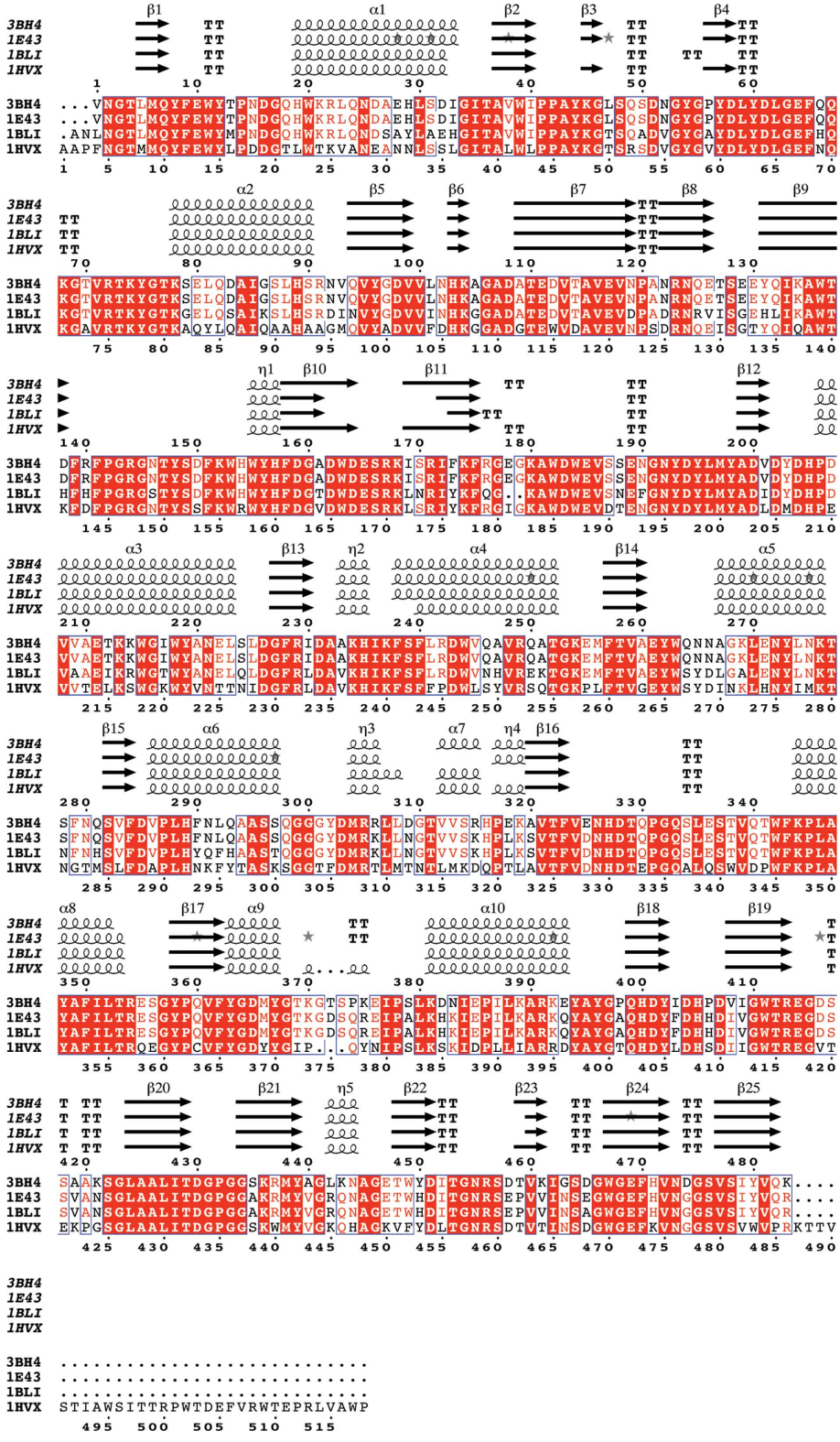

Figure 2
Sequence alignment of BAA, BA2, BLA and BStA calculated with the program *ClustalW*. Secondary structures are presented above the alignment. α-Helices are represented by α, β-strands by β and 3_{10} helices by η. Conserved residues are specified by a red background and similar residues that belong to the same family are shown as red letters.

domain containing residues 1–101 and 203–392 that folds into an (α/β)$_8$-barrel and has slight discrepancies in secondary-structural elements in regions 44–58 and 304–320 among the structures discussed here. In this stable domain, eight helices with variable length encompass eight strands of different sizes. The high-resolution crystal structure of BAA revealed that as in BA2 and BStA there are two extra short β-strands (β3 and β4) located between Aα1 and Aα2 which are absent in BLA (Fig. 2). In addition, the absences of α7 and η4 in BA2 and BLA, respectively, are surprising because of their identical sequence with respect to BAA. The active site and conserved Ca^{2+} sites are located in the C-terminal region of the central domain A (Machius *et al.*, 1995). Another interesting feature in this domain is the presence of a π-helix in the region 370–373 in BStA that it is absent in the other structures. The C-terminal portion of the protein (residues 393–484) folds into a Greek-key motif. This all-β domain contains eight strands and one 3_{10}-helix located between Cβ4 and Cβ5. It is conserved in nearly all α-amylases and its functional role has not been recognized (Kadziola *et al.*, 1994). The only difference in this domain is the shorter length of β23 (Cβ6) in BA2 and BLA than in BAA and BStA (Fig. 2). The most variable domain in α-amylases, domain B, is an excursion from domain A located between Aα3 and Aβ3. This domain is absent in saccharifying α-amylases such as *B. subtilis* α-amylase (BSuA), while in thermostable bacterial α-amylases it has striking features and is elongated by up to 100 residues. In this domain, Bβ1 and Bβ3 wrap around each other and are twisted and loosely connected to four other shorter β-sheets. Sequence alignment in this domain indicated that BLA and BA2 have shorter β10 and β11 strands in the region 155–175 relative to BAA and BStA. The high-resolution crystal structure of BAA roughly indicated greater flexibility in this domain relative to BLA as evaluated from temperature factors (Fig. 1*a*). However, the values of crystallographic thermal parameters are not characteristic of a particular protein and should therefore be treated with caution considering their high dependence on the mode of refinement (artificially low *B* factors can be obtained upon over-refinement) and molecular packing in the crystal. This flexibility would be expected since BAA is far less thermostable than BLA and indeed thermal denaturation starts from this domain. It appears that BAA resembles BA2 and BStA more closely than BLA in terms of its secondary-

Table 3
The residues involved in Ca coordination of monomer *A* and their corresponding distance from the metal ions (monomer *B* in the asymmetric unit has the same properties).

Metal ion	Ligand	Distance (Å)
Calcium I	Asp201 OD1	2.5
	Asp195 O	2.5
	Asp195 OD1	2.5
	Asn103 OD1	2.4
	Wat486	2.7
Calcium II	Ala182 O	2.5
	Asp205 OD2	2.6
	Asp203 OD2	2.5
	Asp160 OD1	2.6
	Asp184 OD1	2.5
	Wat566	3.0
Calcium III	Asp431 OD1	2.6
	Asp431 OD2	2.6
	Gly301 O	2.5
	Pro407 O	2.5
	Asp408 OD2	2.4
	Wat305	2.8
Calcium IV	Asn445 OD1	2.5
	Glu448 OE1	2.7
	Glu448 OE2	2.5
	Gln483 OE1	2.6
	Wat500	2.7
	Wat524	2.7
	Wat539	2.9
	Wat550	2.9
Sodium	Asp184 OD2	2.5
	Asp160 OD2	2.4
	Asp201 OD2	2.4
	Asp195 OD2	2.5
	Val202 O	2.5

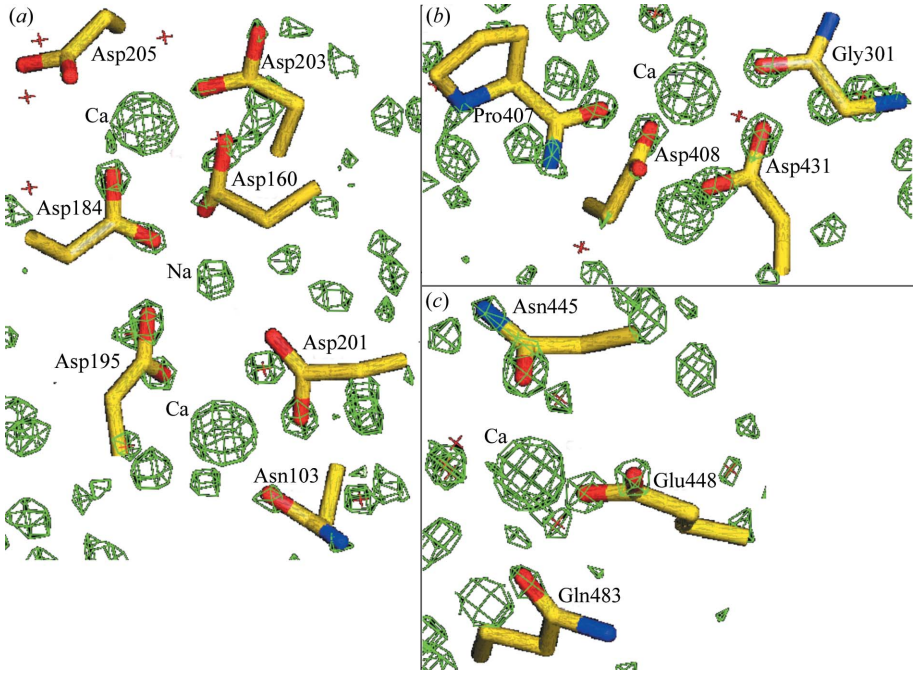

Figure 3
Difference Fourier electron-density map contoured at 2.5σ at metal-binding sites after refinement of the entire model excluding the metal ions. The assigned densities for ions are labelled. (*a*) Ca–Na–Ca metal-binding site. (*b*) Third calcium-binding site at the interface of domains C and A. (*c*) Fourth calcium-binding site in domain C.

structural elements. Accordingly, it has been proved experimentally that BA2 mimics its parental biochemical properties and is an intermediate between BLA and BAA (Brzozowski *et al.*, 2000; Conrad *et al.*, 1995).

3.2. Molecular packing in the crystal

The oligomeric states of functional macromolecular complexes in solution are important for creating an understanding of their functions. In the crystal structure of BAA there are two molecules per asymmetric unit (Fig. 1*b*), which is a notably rare phenomenon in α-amylases. The only polar interactions between the two monomers occur at OD1 and ND2 of Asn222 in chain *A*, which form hydrogen bonds to the main-chain N atom of Ala421 and to the carbonyl group of Ser419 in chain *B*, respectively (Fig. 1*b*). Although the crystal structures of BAA obtained in three other conditions were consistent with a dimer, size-exclusion chromatography and dynamic light-scattering data did not confirm the presence of a dimeric form of BAA (data not shown). As there is no evidence in the literature for such a dimeric form of BAA in solution (Takkinen *et al.*, 1983; Walker & Campbell, 1967*a*,*b*; Borgia & Campbell, 1978), this phenomenon could be assigned to either the crystal packing of BAA or salt-induced oligomerization (Vielle & Zeikus, 2001).

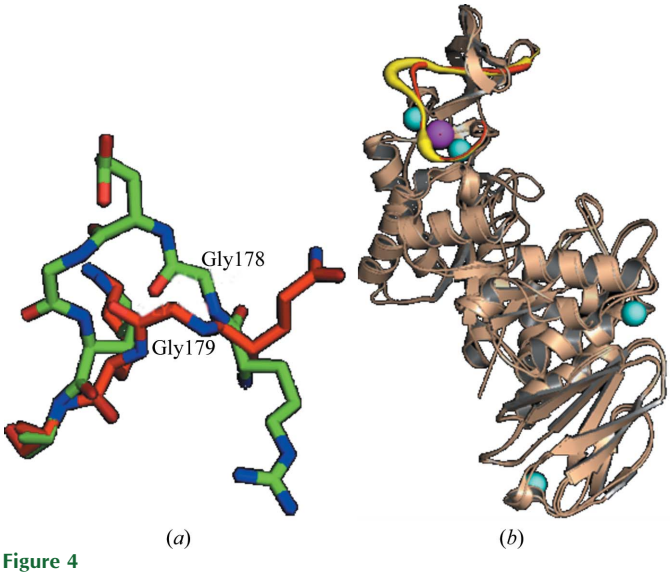

Figure 4
Superposition of BAA (green) and BLA (red). (*a*) The bulge created by two extra residues in BAA and main-chain deviation from the critical Gly178. (*b*) Superimposition of the BAA (yellow) and BLA (red) structures; the thickness of the coils represents their isotropic *B* factors in the TSD3 region. Calcium and sodium ions are presented as cyan and magenta spheres, respectively.

3.3. Determination of metal ions in the structure

Metal ions were identified crystallographically by decreasing the contour levels in both $|F_o - F_c|$ and $|2F_o - F_c|$ maps until the electron density of water molecules vanished. The positions of the ion-binding sites were confirmed by both alignment of the sequence of BAA with those of other homologues (BLA and BStA) and the nature of the atoms surrounding the cations. These maxima also resembled a distorted pentagonal bipyramid, which is a stereochemically rational coordination geometry for calcium. We determined the existence of four Ca^{2+} ions in each monomer, two of which are located in domain B 8.6 Å apart from each other, with one sodium ion between them at a distance of 4 Å from Ca^{2+} I (Fig. 3*a*). The third Ca^{2+} ion lies at the interface of domains A and C (Fig. 3*b*) and the fourth Ca^{2+} ion lies in domain C (Fig. 3*c*). Table 3 lists residues that interact with cations and their corresponding distances.

The known α-amylase structures discussed here share a common calcium-binding site located at the interface of domains A and B in close proximity to the active site (Fig. 4). This site seems to be essential for structural integrity and catalytic activity of the protein (Janecek, 1997) and containes two calcium ions that are 8.6 Å apart from each other and that are bridged by a sodium ion at a distance of 4 Å from Ca^{2+} I. This site has been experimentally verified to play an important role in the thermostability of BLA (Machius *et al.*, 1998; Declerck *et al.*, 2000). It has been suggested that these metal centres are highly sequence-specific and are only found in thermostable homologues (Linden & Wilmanns, 2004). In the present study, the presence of the sodium ion in BAA was proved by the shape of the

Residues	2–10	11–118	119–131	132–161	162–176	177–185	186–194	195–245	246–255	256–313	314–321	322–369	370–384	385–414	415–423	424–455	456–484
2–10	1.46	1.30	1.43	1.50	1.92	8.92	3.12	1.15	1.14	1.08	1.61	1.11	1.40	1.49	4.33	1.42	1.65
11–118	1.3	1.34	1.34	1.28	1.39	3.96	1.72	1.14	1.31	1.13	1.31	1.20	1.29	1.22	2.71	1.28	1.39
119–131	1.43	1.34	1.70	1.74	1.82	7.69	2.84	1.17	1.46	1.26	1.74	1.29	1.72	1.29	3.72	1.45	1.77
132–161	1.5	1.28	1.74	1.39	1.55	5.78	2.49	1.16	1.49	1.16	1.54	1.20	1.43	1.22	3.45	1.22	1.56
162–176	1.92	1.39	1.82	1.55	1.76	6.95	3.07	1.24	1.57	1.33	1.87	1.33	1.91	1.35	3.73	1.59	1.94
177–185	8.92	3.96	7.69	5.78	6.95	11.12	9.07	5.58	8.72	5.33	9.39	5.45	7.45	6.36	9.26	6.41	6.65
186–194	3.12	1.72	2.84	2.49	3.07	9.07	4.22	2.02	3.01	2.02	3.19	2.09	2.93	2.29	4.90	2.32	2.55
195–245	1.15	1.14	1.17	1.16	1.24	5.58	2.02	0.95	1.12	0.98	1.06	1.02	1.16	0.98	3.20	1.00	1.28
246–255	1.14	1.31	1.46	1.49	1.57	8.72	3.01	1.12	1.15	1.13	1.43	1.06	1.35	1.04	4.72	1.24	1.56
256–313	1.08	1.13	1.26	1.16	1.33	5.33	2.02	0.98	1.13	0.99	1.29	1.01	1.08	0.92	3.21	1.11	1.35
314–321	1.61	1.31	1.74	1.54	1.87	9.39	3.19	1.06	1.43	1.29	1.73	1.21	1.59	1.32	4.17	1.51	1.74
322–369	1.11	1.20	1.29	1.20	1.33	5.45	2.09	1.02	1.06	1.01	1.21	0.98	1.20	0.90	2.96	1.01	1.24
370–384	1.4	1.29	1.72	1.43	1.91	7.45	2.93	1.16	1.35	1.08	1.59	1.20	1.47	1.22	3.74	1.39	1.57
385–414	1.49	1.22	1.29	1.22	1.35	6.36	2.29	0.89	1.04	0.92	1.32	0.90	1.22	0.85	4.05	0.94	1.24
415–423	4.33	2.71	3.72	3.45	3.73	9.26	4.90	3.20	4.72	3.21	4.17	2.96	3.74	4.05	5.91	3.27	3.39
424–455	1.42	1.28	1.45	1.22	1.59	6.41	2.32	1.00	1.24	1.11	1.51	1.01	1.39	0.94	3.27	1.22	1.37
456–484	1.65	1.39	1.77	1.56	1.94	6.66	2.55	1.28	1.56	1.35	1.74	1.24	1.57	1.24	3.39	1.37	1.83

Figure 5
The *TLSMD* optimization algorithm matrix for BAA as sequential segments. The main-chain anisotropic ADPs for the region of residues 177–185, a loop located at the surface of domain B, differs considerably from the rest, implying greater internal flexibility. Loops with greater flexibility are shaded in grey.

electron density, by sequence comparisons with other homologues and by consideration of the crystallization conditions, although it does not warrant further experimental confirmation. There are no major differences in this critical metal-binding loop compared with other homologues (BLA, BA2 and BstA) and other areas of the structure should be inspected in order to investigate the thermostability, although the bulge created in this region in BAA (Fig. 4*a*) could facilitate displacement of the loop and the diffusion of cations when the temperature rises. Deletion of Arg-Gly in this region could facilitate thermostabilization of BLA (Suzuki *et al.*, 1989; Janecek, 1997; Machius *et al.*, 1998).

3.4. Analysis of domain displacement in the crystal and thermostability

A translation/libration/screw motion determination (*TLSMD*) optimization algorithm model (Painter & Merritt, 2006) of BAA as sequential protein segments shows the r.m.s.d. *B* values of the individual groups on the diagonal and the r.m.s.d. *B* values of combined groups as off-diagonal elements (Fig. 5). The C-terminal and central domain of BAA indicated relatively small and uniform differences values for the main-chain atoms, suggesting that there is little relative displacement in these domains, with the exception of a segment comprising residues 415–423, which is a loop located at the surface of domain C. In contrast, two consecutive groups in domain B involving residues 177–194 showed elevated difference values, especially in the region 177–185, indicating that it contains loops with a substantial amount of internal flexibility (grey parts in Fig. 5). It is noteworthy that this loop in domain B constructs part of the cage that is responsible for calcium binding.

Temperature exerts a profound influence on the balanced interplay of structural flexibility and rigidity. A comparison of protein flexibility between homologous proteins was undertaken by measuring the structural dynamics in various ways; amide-proton exchange kinetics and neutron spectroscopic studies of BAA and BLA indicated that BLA had a more flexible structure (Fitter, 2005). Visual inspection of the isotropic temperature factors of the crystal structures of BAA and BLA at high resolution indicated higher *B* factors for BAA in domain B (Fig. 1*a*). However BLA, in contrast to BAA, showed increased *B* factors in domain A especially at Aα1, Aα3 and the C-terminal side of Aα5. It has been definitely determined that domain B plays a paramount role in the thermal stability of liquefying-type α-amylases (Machius *et al.*, 1995, 1998; Declerck *et al.*, 2000) and in fact thermal denaturation is triggered from this section. Therefore, an increase in the rigidity of domain B would have a profound effect on the whole stability by delaying triggering of the unfolding process. Nevertheless, since BAA and BLA unfold irreversibly when subjected to high temperatures, the above possibilities may only account for reversible steps occurring just before the irreversible step and will remain speculative until strong experimental data lead to a conclusion on this issue.

3.5. Structural analysis, accessible surface area and electrostatic potential charge

Thermostability of a protein is generally achieved through subtle interactions of many types. Despite BAA and BLA having more than 80% homology, the latter has a half-life that is more than 100 times greater than the former at 363 K and is stable and active at high temperatures (Tomazic & Klibanov, 1988). Although much effort has been undertaken to elucidate the basis of this difference, the structural features leading to thermostability remain controversial, partly owing to the lack of a crystal structure of BAA for comparison. To this end, we have determined the crystal structure of BAA at high resolution and examined the region previously proposed to be responsible for the greater stability of BLA. Conrad *et al.* (1995) constructed chimeric variants of α-amylase using gene shuffling and proposed thermostability determinants (TSDs) according to which BLA was found to be more stable. The first region (TSD1) composed of residues 34–74 played no significant role in the thermostability, but was crucial in terms of alterations in the content of maltohexaoside and maltopentaoside in the digest. As expected from its importance in substrate specificity, TSD1 was almost identical in BAA and BLA and its superimposition indicated no significant r.m.s.d. for the main chain in both. TSD2 covers residues 112–142, in which there are five intra-side-chain interactions in BAA and three in BLA. In BAA, Arg141 NE forms salt bridges to OD1 and OD2 of Asp139, whereas these interactions are lost in BLA because both corresponding positions are replaced by histidines. Another important hydrogen bond found in BAA is formed between Tyr132 OH and Glu118 OE2 and is absent from BLA because Tyr132 is replaced by His. BLA is a hyperthermophilic counterpart and thus a quantitatively larger number of polar interactions would be expected in BLA relative to BAA. This controversial point might be explained by a detailed inspection of the amino-acid compositions of these amylases. BLA has ten more histidines, which are spread out over the whole protein, compared with BAA, in which seven histidines are replaced by polar residues. Basic principles make it obvious that not only the folded state but also the unfolded state has an impact on protein stability and on the unfolding transition (Dao-pin *et al.*, 1991). This property is

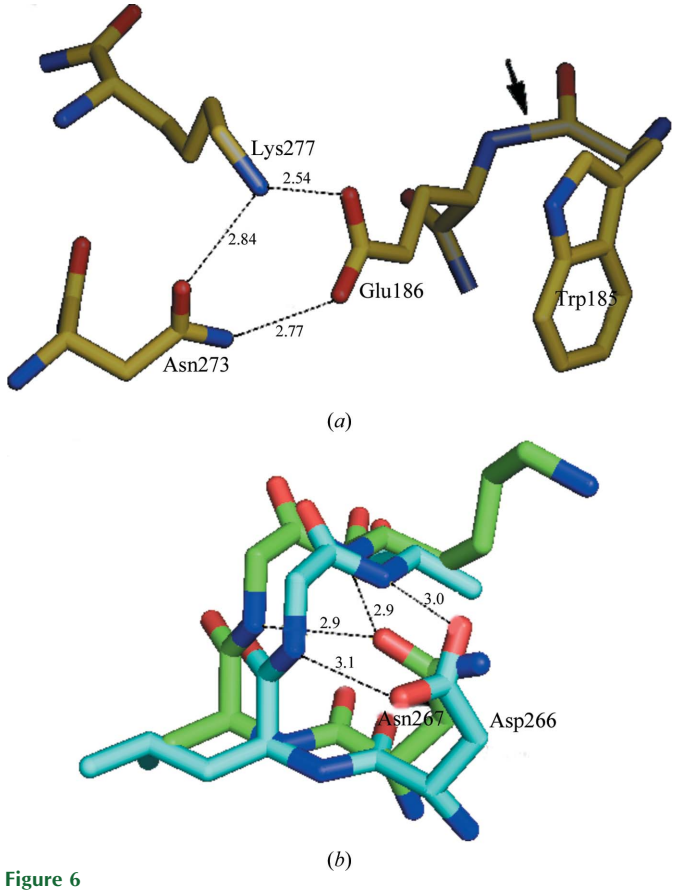

Figure 6
The *cis*-peptide bond stabilized by the interaction of Glu186 with Lys277 and Asn273. (*a*) The *cis*-peptide bond is shown by an arrow. (*b*) Differences in the side-chain conformation and hydrogen-bonding features of Asn in BAA (green) and Asp in the same position in BLA (cyan). Distances are specified in Å.

more or less directly related to the conformational entropy of the unfolded state; the native protein exhibits a rather restricted conformational freedom and upon unfolding a larger degree of conformational freedom is accompanied by a less compact structure. Because the hydration of nonpolar groups apparently destabilizes proteins (Dao-pin *et al.*, 1991), such an increase in the number of nonpolar residues in the sequence of BLA is suggested to contribute to the stabilization of its native state through decreased entropy of the unfolded state upon thermal denaturation. A similar mechanism, in part rigidifying, was applied to obtain a hyperthermostable enzyme by decreased entropy of the unfolded state *via* the introduction of Gly→Ala and Ala→Pro substitutions (Van Den Burg *et al.*, 1998). However, mutational experiments involving the introduction of His residues into BAA and examination of differences in thermostability would corroborate the above hypothesis.

The third region (TSD3) is constituted of residues 174–179 as proposed by Conrad *et al.* (1995) and overlaps with region I (177–186) reported by Suzuki *et al.* (1989). Located in domain B, this region encompasses metal-binding sites by forming part of the cage trapping a Ca–Na–Ca triad (Machius *et al.*, 1995, 1998; Declerck *et al.*, 2000). Comparison of the crystal structures of BLA and BAA revealed a remarkable main-chain deviation and conformational changes in this loop (Fig. 4*a*). As expected, analysis of the temperature factor of this region exhibited greater fluctuations in BAA (Fig. 4*b*) and hence at high temperatures this loop is likely to be displaced and to let cations diffuse from the structure. Furthermore, the insertion of two residues in BAA and BStA relative to BLA leaves a sufficient cavity inside the structure, whereas in the latter no cavity with a significant probe-accessible volume was found with *VOIDOO* using a probe radius of 1.4 Å (Kleywegt & Jones, 1994; Suvd *et al.*, 2001).

The metal-binding cage of the triad, like those of other α-amylases (BLA, BStA and BA2) of liquefying type, has a notable *cis*-peptide bond between Trp185 and Glu186 (Fig. 6*a*), which has been suggested to be necessary for the orientation of adjacent residues to coordinate properly with the calcium ion (Machius *et al.*, 1998; Declerck *et al.*, 2000). Upon close inspection of the structure, this *cis*-peptide-bond conformation is found to be stabilized by both a salt bridge between Glu186 and Lys277 and the formation of a hydrogen bond between Glu186 and Asn273 from Aα3 (Fig. 6*a*). Such a role for ionic interactions and hydrogen bonding in thermostabilization of α-amylases has been considered previously by mutational studies (Kim *et al.*, 2003; Declerck *et al.*, 1997, 2000, 2003).

Superposition of the final section (TSD4) affecting thermostability neither showed a main-chain distortion nor side-chain conformational changes. Nevertheless, it cross-talks with TSD3 as it includes Lys277 and Asn273, which stabilize a *cis*-peptide-bond conformation *via* ionic interactions with Glu186. Another distinction of BAA from BLA is that Asn266 in BAA is replaced by Asp in BLA (Yuuki *et al.*, 1985), which has a weaker role in helix stabilization, since in contrast to Asp it can hydrogen bond to the main chain only through its OD1 (Fig. 6*b*). Despite the high degree of similarity between BAA and the other amylases, we have found regions of BAA in which the propensity of the constituent amino acids to adopt a specific secondary structure differs as evaluated by *ESPript* (Gouet *et al.*, 1999). For example, the region consisting of residues 161–172, which contains a loop located at the surface of domain B, differs in length in different amylases. This loop has the shortest length in BAA and BStA, whereas it has the longest length in BLA. The calculated accessible surface areas (ASAs) of these structures did not imply much difference (less than 1%) and hence statistically there is no significant positive correlation between accessible surface areas (ASAs) and thermostability at least for this set of proteins (Table 2). Such a view has been maintained in a comparative survey of α-amylases from different origins (Linden & Wilmanns, 2004; Aghajari *et al.*, 1998). In contrast to these results, the total electrostatic potential charges indicate greater variation between these structures (Table 2).

4. Conclusion

In summary, the crystallization and solution of the crystal structure of BAA at high resolution enabled us to pinpoint that there are factors exclusive to BLA that make it more thermostable compared with BAA, *e.g.* notable loop stabilization occurring mainly through the entropic effect of *X*→His substitutions, a shortening of the loop responsible for the calcium-binding site and subtle intramolecular interactions that both decrease its freedom for conformational changes and link juxtaposed portions of the structure and improve the anchoring to the rest of the protein. Moreover, *TLSMD* analysis of anisotropic atomic displacement parameters (ADPs) indicated that domain B in BAA behaves as a more flexible part and indeed the dynamic properties of this domain have a profound impact on the function and stability of the protein. As the challenge is to address which factors are critical for the stability of a specific protein, among the various determinants of structural stability under extreme conditions we emphasize the importance of entropic effects of the native or unfolded state in thermal stabilization as well as the deleterious effect of loop flexibility in the part of the protein which is thought to trigger the unfolding process of α-amylases. The structure provided in this study could be used for further comparative structural analysis correlating the structure–activity relationship.

We thank the support staff for their technical assistance and discussion during data collection on beamlines BL13B1 and13C1 of NSRRC in Taiwan. The authors express their gratitude to the research council of Tarbiat Modares University and the National Synchrotron Radiation Research Center (NSRRC) and the National Science Council (NSC) in Taiwan for financial support (NSRRC 963RSB02, NSC 95-2311-B213-001-MY3, NSC 95-2923-B-213-001-MY3 to C-JC) during the course of this project.

References

Aghajari, N., Feller, G., Gerday, C. & Haser, R. (1998). *Structure*, **6**, 1503–1516.
Borgia, P. T. & Campbell, L. L. (1978). *J. Bacteriol.* **134**, 389–393.
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.
Brzozowski, A. M., Lawson, D. M., Turkenburg, J. P., Bisgaard-Frantzen, H., Svendsen, A., Borchert, T. V., Dauter, Z., Wilson, K. S. & Davies, G. J. (2000). *Biochemistry*, **33**, 9099–9107.
Chen, M. E., Cang, H. X. & Nymeyer, H. (2007). *NOC* v.3.01. http://noch.sourceforge.net.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Conrad, B., Hoang, V., Polley, A. & Hofemeister, J. (1995). *Eur. J. Biochem.* **230**, 481–490.
Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W. & Matthews, B. W. (1991). *Biochemistry*, **30**, 11521–11529.
Declerck, N., Machius, M., Chambert, R., Wiegand, G., Huber, R. & Gaillardin, C. (1997). *Protein Eng.* **10**, 541–549.
Declerck, N., Machius, M., Joyet, P., Wiegand, G., Huber, R. & Gaillardin, C. (2003). *Protein Eng.* **16**, 287–293.
Declerck, N., Machius, M., Wiegand, G., Huber, R. & Gaillardin, C. (2000). *J. Mol. Biol.* **301**, 1041–1057.
DeLano, W. L. (2002). *The PyMOL Molecular Viewer.* http://www.pymol.org.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Fitter, J. (2005). *Cell. Mol. Life Sci.* **62**, 1925–1937.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Holm, L., Koivula, A. K., Lehtovaara, P. M., Hemminki, A. & Knowles, J. K. C. (1990). *Protein Eng.* **3**, 181–191.
Ihara, H., Sasaki, T., Tsuboi, A., Yamagata, H., Tsukagoshi, N. & Udaka, S. (1985). *J. Biochem.* **98**, 95–103.
Janecek, S. (1997). *Prog. Biophys. Mol. Biol.* **67**, 67–97.
Kadziola, A., Abe, J., Svensson, B. & Haser, R. (1994). *J. Mol. Biol.* **239**, 104–121.
Kleywegt, G. J. & Jones, T. A. (1994). *Acta Cryst.* D**50**, 178–185.
Kim, Y.-W., Choi, J.-H., Kim, J.-W., Park, C., Kim, J.-W., Cha, H., Lee, S.-B., Oh, B.-H., Moon, T.-W. & Park, K.-H. (2003). *Appl. Environ. Microbiol.* **69**, 4866–4874.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
Linden, A. & Wilmanns, M. (2004). *Chembiochem*, **5**, 231–239.
Machius, M., Declerck, N., Huber, R. & Wiegand, G. (1998). *Structure*, **6**, 281–292.
Machius, M., Wiegand, G. & Huber, R. (1995). *J. Mol. Biol.* **264**, 545–559.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nagashima, T., Tada, S., Kitamoto, K., Gomi, K., Kumagai, C. & Toda, S. (1992). *Biosci. Biotechnol. Biochem.* **56**, 207–210.
Nakajima, R., Imanaka, T. & Aiba, S. (1985). *J. Bacteriol.* **163**, 401–406.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* D**62**, 439–450.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.
Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 73–82.
Smirnova, N. A., Sorokin, A. V., Lapdev, D. A., Veiko, V. P. & Kozlov, Y. I. (1987). *Mol. Biol. (Mosk.)*, **22**, 1257–1271.
Suvd, D., Fujimoto, Z., Takase, K., Matsumura, M. & Mizuno, H. (2001). *J. Biochem.* **129**, 461–468.
Suzuki, A., Yamane, T., Ito, Y., Nishio, T., Fujiwara, H. & Ashida, T. (1990). *J. Biochem.* **108**, 379–381.
Suzuki, Y., Ito, N., Yuuki, T., Yamagata, H. & Udaka, S. (1989). *J. Biol. Chem.* **264**, 18933–18938.
Svensson, B. (1991). *J. Jpn. Soc. Starch Sci.* **38**, 125–135.
Svensson, B. & Sogaard, M. (1992). *Biochem. Soc. Trans.* **20**, 34–41.
Takase, K., Matsumoto, T., Mizuno, H. & Yamane, K. (1992). *Biochim. Biophys. Acta*, **1120**, 281–288.
Takkinen, K., Pettersson, R. F., Kalkkinen, N., Palva, I., Soderlund, H. & Kaariainen, L. (1983). *J. Biol. Chem.* **258**, 1007–1013.
Thompson, J. D., Higgins, D. J. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Tomazic, S. J. & Klibanov, A. M. (1988). *J. Biol. Chem.* **263**, 3092–3096.
Van Den Burg, B., Vriend, G., Veltman, O. R., Venema, G. & Eijsink, V. G. H. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2056–2060.
Vielle, C. & Zeikus, J. G. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
Vihinen, M., Ollikka, P., Niskanen, J., Meyer, P., Suominen, I., Karp, M., Holm, L., Knowles, J. & Mantsala, P. (1990). *J. Biochem.* **107**, 267–272.
Walker, N. E. & Campbell, L. L. (1967*a*). *J. Bacteriol.* **94**, 1131–1135.
Walker, N. E. & Campbell, L. L. (1967*b*). *Biochemistry*, **6**, 3681–3689.
Yamazaki, H., Ohmura, K., Nakamura, A., Takeichi, Y., Otozai, K., Yamazaki, M., Tamura, G. & Yamane, K. (1983). *J. Bacteriol.* **156**, 327–337.
Yang, M., Gallizzi, A. & Henner, D. (1983). *Nucleic Acids Res.* **11**, 237–249.
Yuuki, T., Nomura, T., Tezuka, H., Tsuboi, A., Yamagata, H., Tsukagoshi, N. & Udaka, S. (1985). *J. Biochem.* **98**, 1147–1156.