doi:10.1016/S0022-2836(03)00630-2 J. Mol. Biol. (2003) 330, 607–620

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Crystal Structure of a Natural Circularly Permuted Jellyroll Protein: 1,3-1,4-β-D-Glucanase from *Fibrobacter succinogenes*

Li-Chu Tsai[1,2], Lie-Fen Shyur[3], Shu-Hua Lee[3], Su-Shiang Lin[3] and Hanna S. Yuan[1]*

[1]*Institute of Molecular Biology Academia Sinica, Taipei Taiwan, ROC*

[2]*Department of Molecular Science and Engineering National Taipei University of Technology, Taipei, Taiwan ROC*

[3]*Institute of BioAgricultural Sciences, Academia Sinica Taipei, Taiwan, ROC*

**Corresponding author*

The 1,3-1,4-β-D-glucanase from *Fibrobacter succinogenes* (Fsβ-glucanase) is classified as one of the family 16 glycosyl hydrolases. It hydrolyzes the glycosidic bond in the mixed-linked glucans containing β-1,3- and β-1,4-glycosidic linkages. We constructed a truncated form of recombinant Fsβ-glucanase containing the catalytic domain from amino acid residues 1–258, which exhibited a higher thermal stability and enzymatic activity than the full-length enzyme. The crystal structure of the truncated Fsβ-glucanase was solved at a resolution of 1.7 Å by the multiple wavelength anomalous dispersion (MAD) method using the anomalous signals from the seleno-methionine-labeled protein. The overall topology of the truncated Fsβ-glucanase consists mainly of two eight-stranded anti-parallel β-sheets arranged in a jellyroll β-sandwich, similar to the fold of many glycosyl hydrolases and carbohydrate-binding modules. Sequence comparison with other bacterial glucanases showed that Fsβ-glucanase is the only naturally occurring circularly permuted β-glucanase with reversed sequences. Structural comparison shows that the engineered circular-permuted *Bacillus* enzymes are more similar to their parent enzymes with which they share ~70% sequence identity, than to the naturally occurring Fsβ-glucanase of similar topology with 30% identity. This result suggests that protein structure relies more on sequence identity than topology. The high-resolution structure of Fsβ-glucanase provides a structural rationale for the different activities obtained from a series of mutant glucanases and a basis for the development of engineered enzymes with increased activity and structural stability.

Keywords: glycosyl hydrolases; circular permutation; Ca^{2+} binding; clan GH-B; family 16 GH

Introduction

Circular permutations in protein structures occur while the original N and C-terminal regions of one protein are linked to form a continuous part of polypeptide chain and new termini are formed elsewhere. Based on sequence alignment, circular permutation can be identified in a protein showing that its N-terminal sequence is similar to the C-terminal sequence of the other protein and *vice versa*.[1] The first observation of a naturally occurring circular permutation was described about two decades ago in amino acid sequences of favin *versus* concanavalin A.[2] Since then naturally occurring circular permutations have been found in several protein families,[3] including bacterial β-glucanases,[4] FMN-binding proteins,[5] DNA adenine-N6 methyltransferases,[6] β-glucosidases,[7] cytosine-C5 methyltransferases,[8] swaposin,[9] α-1,3 and α,-1,6 glucan-synthesizing glucosyltransferases,[10] transaldoase,[11] double-ψ β-barrels,[12] glutathione,[13] ferredoxins[14] and serine proteinase inhibitors.[15] Recently, many more pairs of circularly permuted proteins with identical, related or different

Abbreviations used: Fsβ-glucanase, *Fibrobacter succinogenes* 1,3-1,4-β-D-glucanase; TFsβ-glucanase, truncated form of Fsβ-glucanase; MAD, multiple wavelength anomalous dispersion; rms, root-mean-square.

E-mail address of the corresponding author: hanna@sinica.edu.tw

0022-2836/03/$ - see front matter

functions have been identified in protein databases, based on protein sequence or structure comparisons.[1,16] In contrast to abundant sequence analyses, fewer studies have been conducted at the three-dimensional structural level for naturally occurring circularly permuted proteins. The only known three-dimensional structures of a pair of naturally occurring circularly permuted proteins, with homologous sequences, are the crystal structures of pea lectin[17] and jack bean concanavalin A.[18,19] This is a special case, however, in that the circular permutation occurs as a post-translational modification at protein but not at gene level.[20,21] In the past, the most detailed analyses of circular permutation at the three-dimensional structural level have been carried out using series of engineered *Bacillus* β-glucanase mutants.[22]

β-Glucanases (1,3-1,4-β-D-glucan 4-glucanohydrolases, EC 3.2.1.73) hydrolyze mixed-linked glucans containing β-1,3- and β-1,4-glycosidic linkages, such as β-glucans from grain endosperm cell walls or lichenan from Islandic moss.[23] 1,3-1,4-β-D-Glucanases have been identified from bacteria and plants, including different *Bacillus* species,[24–31] *Fibrobacter succinogenes*,[32] *Ruminococcus flavefaciens*,[33] *Clostridium thermocellum*[34] and barley.[35,36] Bacterial β-glucanases are classified as members of the family 16 glycosyl hydrolases of clan GH-B, which hydrolyze the glycosidic bond based on a general acid/base catalysis with a net retention of the anomeric configuration.[37–39] Crystal structures of several bacterial 1,3-1,4-β-D-glucanases have been reported, including two native enzymes from *Bacillus macerans*[40] and *Bacillus licheniformis*[41] and several genetically engineered enzymes, including the *Bacillus* hybrid enzyme H(A16-M), in which the first 16 N-terminal residues are from *Bacillus amyloliquefaciens* and the rest are from *B. macerans*,[42,43] and the circularly permuted proteins derived from H(A16-M).[4,44] These engineered circularly permuted glucanase enzymes of H(A16-M) show almost identical jellyroll β-sandwich structures to their parent enzyme with rms differences below 0.4 Å.[44] In addition to *Bacillus* β-D-glucanase, the crystal structure of the engineered α-spectrin SH3 domain is the only example to demonstrate the structural consequences resulting from circular permutation.[45] The structure of the engineered circularly permuted SH3 domain is almost identical to that of the wild-type protein with rms deviations of less than 1 Å. Crystal structure analyses of engineered *Bacillus* β-glucanases and SH3 domains reveal that these circularly permuted proteins have native-like structures and lead to the conclusion that the protein structure depends more on sequence identity than topology. However, apart from permutation, natural forms of circularly permuted proteins have more variations in sequences than those of the engineered proteins, therefore, it is necessary to solve the structures of natural circularly permuted proteins in order to confirm the conclusion derived from engineered proteins.

A comparison amongst all bacterial β-glucanases has shown that 1,3-1,4-β-D-glucanase from *F. succinogenes* (Fsβ-glucanase) is the only natural circularly permuted enzyme in which two highly conserved catalytic domains of the enzyme are in a reverse orientation as compared to that of other 1,3-1,4-β-D-glucanases.[22,32,34] A segment of P–X–S–S–S–S was repeated five times and only observed in the C-terminal of the *Fibrobacter* enzyme. Wild-type Fsβ-glucanase consists of 339 amino acid residues, in which residues 1–230 can be aligned with *Bacillus* β-glucanases with ~30% sequence identity when the N-terminal portion (residues 1–171) and the C-terminal portion (residues 175–230) of the Fsβ-glucanase are aligned, respectively, to the C-terminal portion (residues 59–214) and N-terminal portion (residues 1–58) of other bacterial β-glucanases (Figure 1). A truncated form of Fsβ-glucanase (TFsβ-glucanase) with approximately 10 kDa deleted at the C terminus, including the five P–X–S–S–S–S repeated segment, was constructed, which exhibited a higher thermal stability than that of the wild-type full-length enzyme.[46] Kinetic analyses showed that TFsβ-glucanase has a 3.9-fold increase in specific activity and a minor (1.5-fold) decrease in substrate binding affinity resulting in a 2.6-fold increase in overall catalytic efficiency for lichenan relative to the wild-type enzyme. Recently we have identified several amino acid residues in Fsβ-glucanase involved in catalysis and enzyme stability using a site-directed mutagenesis approach.[47,48] Here we report the crystal structure of the truncated form of Fsβ-glucanase, which is the first structural determination of a naturally encoded and circularly permuted protein, as compared to other bacterial glucanases. The structure of Fsβ-glucanase is compared with those of wild-type or genetically engineered *Bacillus* β-glucanases without or with sequence permutation. Our results demonstrate that Fsβ-glucanase, with a permuted sequence, folds into a jellyroll β-sandwich structure similar to that of *Bacillus* enzymes. The three-dimensional structure also provides a structural basis for enzymatic catalysis and an explanation for the different phenotypes exhibited by a series of mutants.

Results and Discussion

Structure determination and overall topology

Full-length and truncated Fsβ-glucanases (residues 1–258) containing the catalytic domain, were both screened for crystallization, however, only the truncated Fsβ-glucanase was crystallized from a PEG4000 solution in the presence of Ca^{2+} in three different crystal forms.[46] Subsequently seleno-methionine incorporated TFsβ-glucanase was expressed and crystallized in orthorhombic space group $P2_12_12_1$ with one molecule per

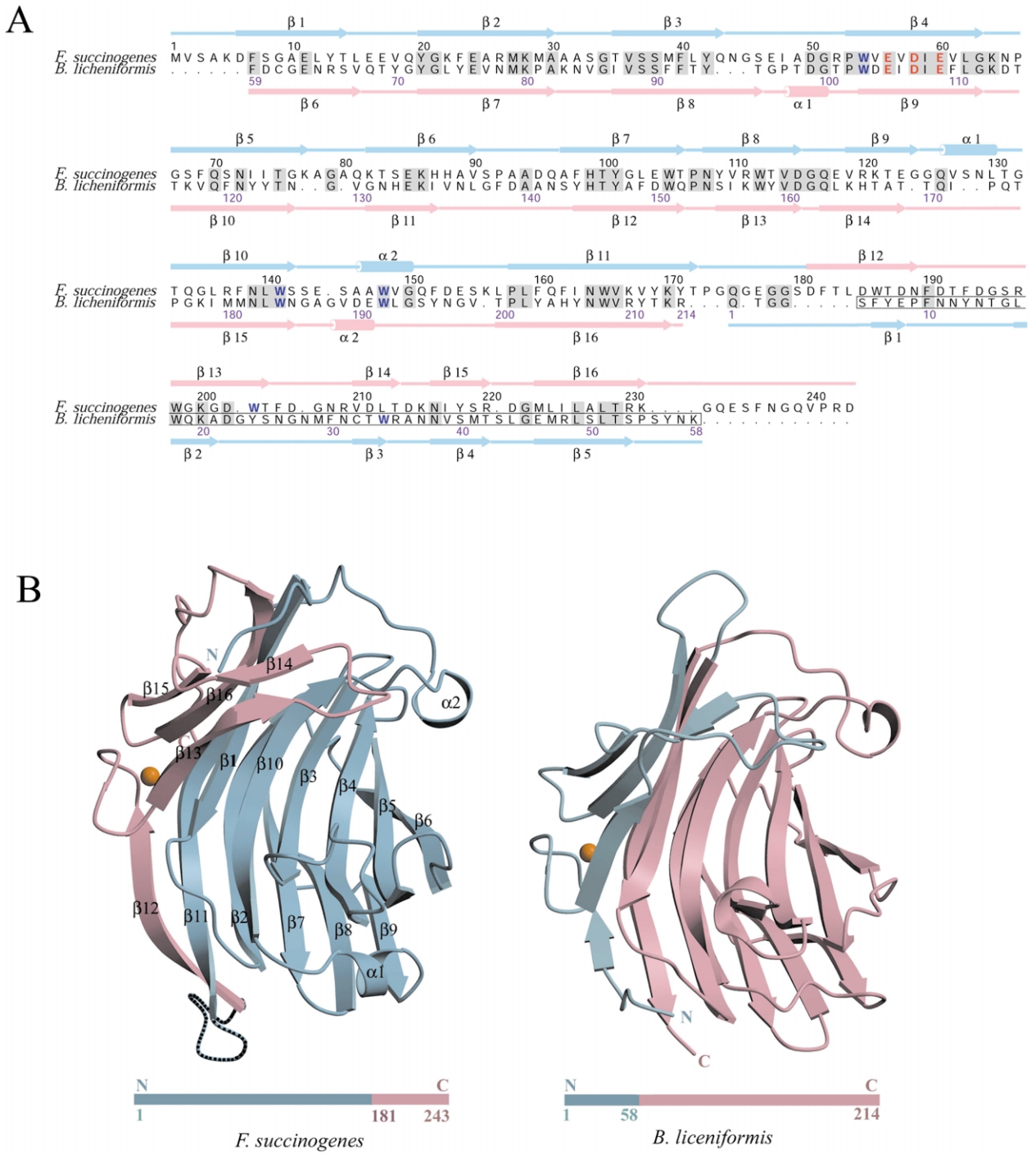

Figure 1. Structure-assisted sequence alignment of 1,3-1,4-β-D-glucanases from *F. succinogenes* (residues 1–243) and *B. licheniformis* using ALSCRIPT.[67] The secondary structural elements identified in the crystal structures, marked in blue or pink according to their location in the N or C-terminal regions, are shown above and below the sequences. The N-terminal peptide of *B. licheniformis* from residues 1 to 58 (framed in the box) has been moved to the C-terminal end to align with the circularly permuted sequence of *F. succinogenes*. The active site residues located on β4 of TFsβ-glucanase are colored red and the tryptophan residues likely involved in substrate binding are colored blue. The gray-shaded residues are identical in the two proteins. (B) The ribbon model of TFsβ-glucanase (the truncated β-glucanase from *F. succinogenes*) in which the C-terminal region (residues 244–339) was deleted; and the wild-type β-glucanase from *B. licheniformis* (PDB ID:1gbg). The N-terminal regions of the two molecules are colored in blue and the C-terminal regions are colored in pink. The N-terminal region of Fsβ-glucanase (residues 1–171) is aligned with the C-terminal region of β-glucanase from *B. licheniformis* (residues 59–214). The C-terminal region of Fsβ-glucanase (residues 175–243) is aligned with the N-terminal region of β-glucanase from *B. licheniformis* (residues 1–58). So these two enzymes represent examples of a pair of naturally occurring and circularly permuted sequences. The calcium ion is displayed as a brown ball located on the convex side of the protein. The loop between β11 and β12 in TFs-glucanase is disconnected in *Bacilli* glucanases to generate N and C termini. The electron density for the loop between β11 and β12 in the *F. succinogenes* is not well defined therefore the loop structure is shown in a dotted line.

Table 1. Data collection, phasing and refinement statistics of the truncated Fsβ-glucanase

Diffraction data statistics			
Wavelength (Å)	$\lambda_1 = 0.9795$ (Se inflection)	$\lambda_2 = 0.9793$ (Se peak)	$\lambda_3 = 0.9600$ (Se remote)
Resolution range (Å)	40–1.7	40–1.7	40–1.7
Observed reflections	107,746	108,457	108,344
Unique reflections	24,394	24,594	24,553
Completeness[a] (%)	97.1 (83.3)	98.0 (92.6)	98.0 (92.8)
$\langle I \rangle / \langle \sigma_I \rangle$[a]	31.6 (5.1)	22.5 (2.9)	22.8 (3.0)
R_{merge} (%)[a,b]	5.5 (19.2)	6.7 (31.1)	6.7 (30.8)
Phasing statistics			
Number of Se sites	4	4	4
Phasing power (centric/acentric)	3.34/3.13	3.04/2.69	1.87/1.66
Figure of merit (centric/acentric)		0.86/0.75	
Refinement statistics			
Resolution range (Å)			40–1.7
Reflections (working/test)			21,494/2343
R_{work}/R_{free} (%)			19.2/23.8
Number of atoms (protein/water/Ca^{2+})			1916/280/1
Average B-factor (Å^2) (protein/water)			16.5/29.3
RMS deviations (bond (Å)/angle (degree))			0.011/1.63

[a] The numbers in parentheses are for the last shell in the resolution range of 1.76–1.70 Å.
[b] $R_{merge} = \sum_h \sum_i |Ih, i - \langle Ih \rangle| / \sum_h \sum_i Ih, i$, where $\langle Ih \rangle$ is the mean intensity of i observations for a given reflection h.

asymmetric unit. The seleno-methionine enzyme exhibits similar catalytic abilities as compared to the TFsβ-glucanase expressed in conventional LB medium (data not shown). The crystal structure was solved by the multiple wavelength anomalous dispersion (MAD) method using the anomalous signals from four selenium atoms at a resolution of 1.7 Å. The Fourier map, calculated using

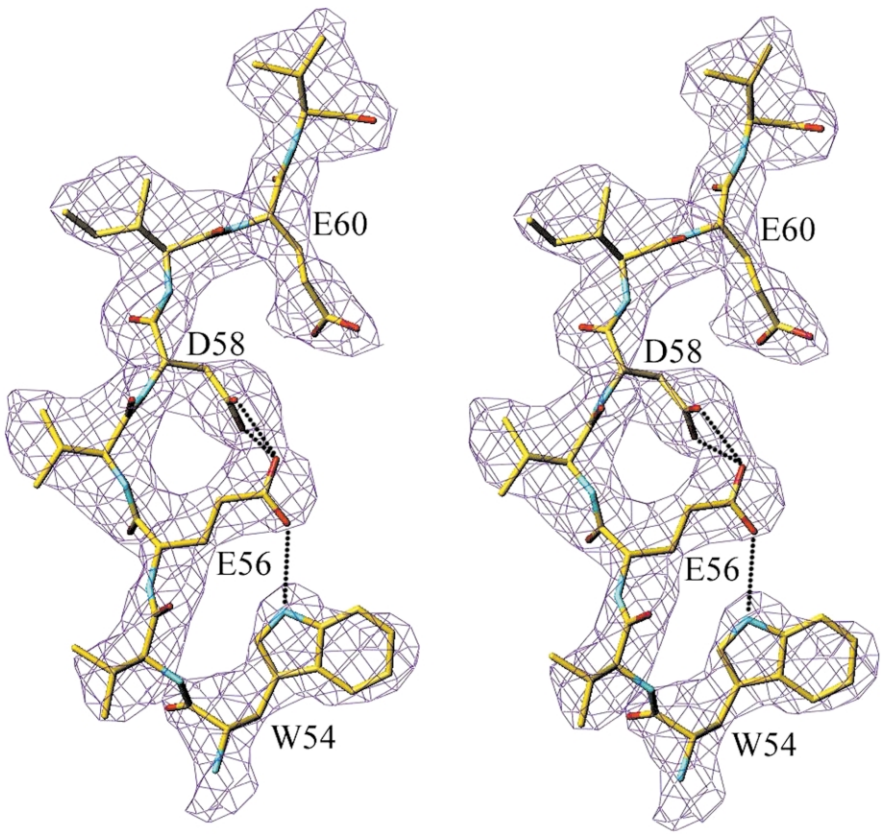

Figure 2. Electron density maps and models demonstrating the quality and resolution of the initial phases. Maps were calculated using solvent-flattened MAD phases at 2.5 Å resolution, contoured at 1σ. A stereo view shows the active site residues of Trp54, Glu56, Asp58 and Glu60 located on the central β4 strand. Dotted lines indicate the hydrogen bonds between residues Trp54/Glu56 and Glu56/Asp58.

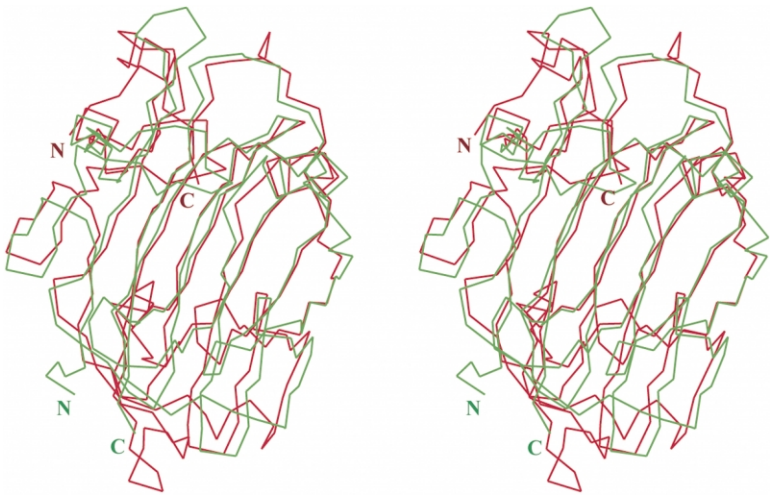

Figure 3. Superimposition of the α-carbon backbones of TFsβ-glucanase (red) and β-glucanases from *B. licheniformis* (PDB ID:1gbg) (green). The average rms difference for the C^{α} backbones of 214 residues between the two structures after least-squares fit is 3.6 Å.

solvent-flattened MAD phases, was of excellent quality (see Figure 2), sufficient for successful structural tracing. The C-terminal residues, 244–258, had no well-defined electron density, therefore the final model consisted of only residues 1–243. The final structure gives an *R*-factor of 21.0% (no σ cutoff) and an *R*-free value of 24.6% (for 9.1% of data) for the data between 40 Å and 1.7 Å (see Table 1). This model shows good stereochemistry with 88.9% residues located in the most favorable region, 10.6% in the generously allowed region and one residue (Tyr20) in the disallowed region (discussed in next section) of the Ramachandran plot calculated by PROCHECK.[49]

The structure of TFsβ-glucanase has a dimension of approximately 50 Å × 45 Å × 30 Å. It comprises 16 β-strands and two α-helices, secondary structural elements shown in Figure 1. The overall fold (Figures 1 and 3) resembles that of a classical sandwich-like β-jellyroll topology mainly consisting of two β-pleated sheets packed back to back with eight β-strands in each sheet. The β-strands are crossed over several times between the two β-sheets, which are surrounded by interconnecting loops and two short α-helices at N-terminal and C-terminal ends of the β-strands. Both β-sheets are twisted and bent, resulting in a convex and a concave side of the molecule. This typical β-jellyroll fold has been observed in a number of glycoside hydrolases, including family 7 endoglucanase I,[50] family 11 xylanases,[51,52] family 12 endoglucanases[50,53,54] and family 16 κ-carrageenase.[55] Several carbohydrate-binding module (CBM) families also bear a similar β-jellyroll fold, including xylan-binding CBM of family 15 (CBM15),[56] CBM4[57] and CBM6.[58] Similar to other glycoside hydrolases with a β-jellyroll structure, the glucan-binding site in Fsβ-glucanase is predicted to be located in a channel of about 28 Å long at the concave side of the molecule.

Structural comparison to *Bacillus* glucanases

A number of crystal structures of *Bacilli* glucanases have been resolved.[44] Here we chose the crystal structure of the wild-type *B. licheniformis* glucanase[41] (PDB entry:1gbg, 1.8 Å resolution) and the engineered circularly permuted form of cpA16M-59[4] (PDB ID:1cpm, 2.0 Å resolution) for the comparison with our naturally occurring circularly permuted TFsβ-glucanase. The overall β-jellyroll folds of the two wild-type glucanases are displayed in Figure 1(B) with the N-terminal regions colored in blue and the C-terminal regions colored in pink. Comparison between ribbon models of TFsβ-glucanase and that of *B. licheniformis* clearly shows that the two proteins start and end at different places but still fold into similar structures. The loop between β11 and β12 (^{173}PGQGEGGS^{180}) in TFsβ-glucanase showing weak electron density with high temperature factor, is disconnected in *Bacillus* glucanases to generate the N and C termini. Similarly, the loop between β5 and β6 in *Bacillus* glucanase, which aligns with β1 and β16 in TFsβ-glucanase, is disconnected in TFsβ-glucanase to generate the N and C termini. The engineered CpHA16M-59 with sequence permutation has a similar connectivity to that of TFsβ-glucanase but its structure is more similar to its parent *Bacillus* enzyme. Based on the sequence alignment showing in Figure 1(A), the least-square fitting was carried out between the engineered *Bacillus* mutant CpHA16M-59 and wild-type bacterial glucanases. Out of 214 residues, the average rms difference for the C^{α} atoms between CpHA16M-59 and *B. licheniformis* glucanase is 0.3 Å; and between CpHA16M-59 and TFsβ-glucanase is 3.6 Å. So the high sequence identity between the two *Bacilli* enzymes (76% after sequence permutation) but not the starting and ending positions of proteins, decides and

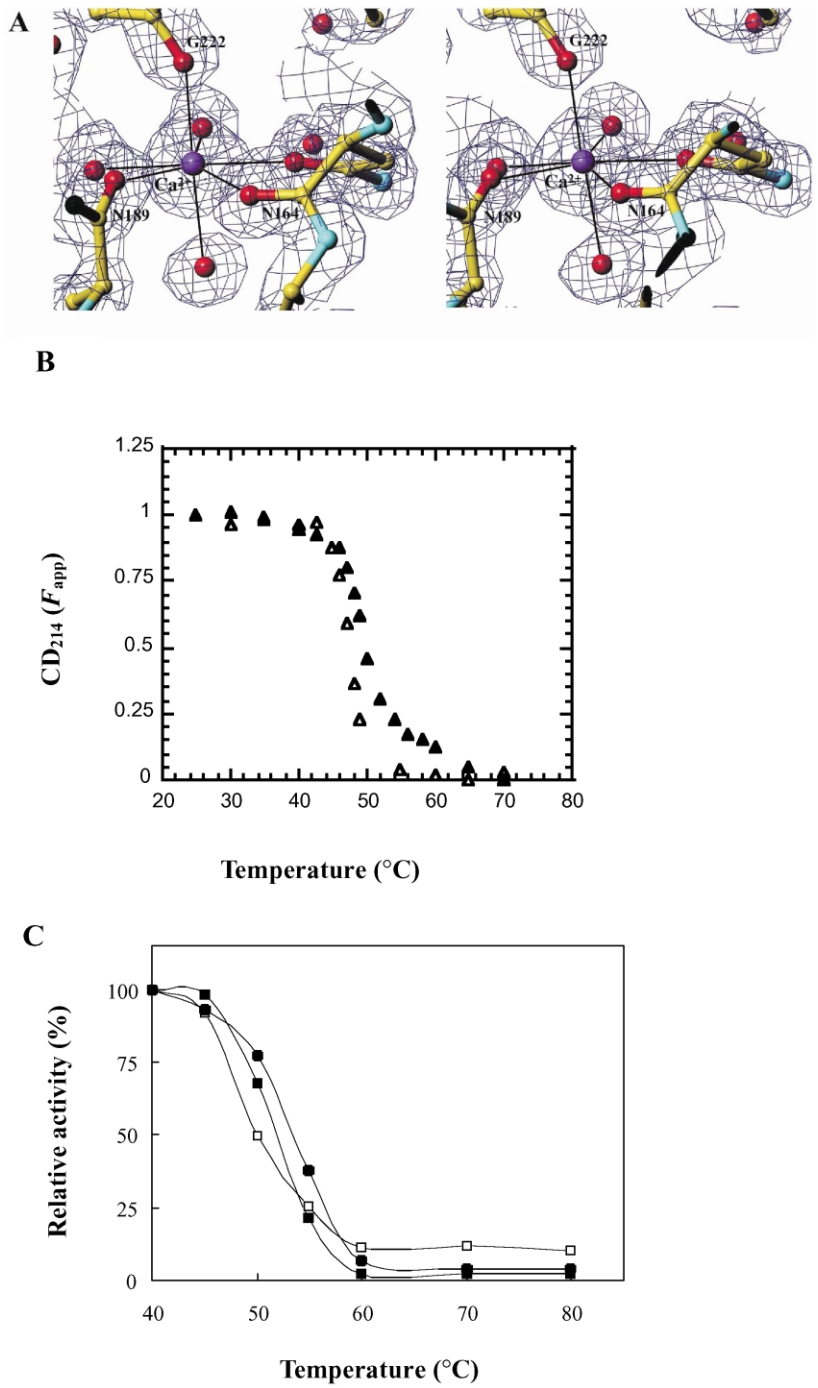

Figure 4. (A) Electron density around the Ca^{2+} binding site of TFsβ-glucanase. Stereo views of the $(2F_o - F_c)$ Fourier map calculated using final model phases contoured at 1.5σ above the mean density at 1.7 Å resolution. (B) A titration curve of the transition of the CD at 214 nm *versus* the apparent fraction of native TFsβ-glucanase at various temperatures, in the presence (▲) or absence (△) of 1 mM Ca^{2+} in 50 mM sodium phosphate buffer (pH 7.0). CD_{214} signals of TFsβ-glucanase at the indicated temperatures were monitored. CD_{214} (F_{app}), representing the apparent fraction of native protein, was calculated as follows: $F_{app} = (Y_{obsd} - Y_U)/(Y_N - Y_U)$.[62] Y_{obsd} represents the observed value of CD at 214 nm of TFsβ-glucanase at various temperatures. Y_N and Y_U represent the CD values at 214 nm of the TFsβ-glucanase at 25 °C and 70 °C, respectively. (C) The effect of Ca^{2+} on heat inactivation of TFsβ-glucanase. TFsβ-glucanase enzyme was pre-incubated in the absence (□) of Ca^{2+}, or in the presence of 1 mM (■) or 50 mM (●) Ca^{2+} in 50 mM sodium citrate buffer (pH 7.0). The residual enzymatic activity after the heat treatment was then immediately measured.

Table 2. Interpretation of kinetic parameters for the mutant forms of *F. succinogenes* 1,3-1,4-β-D-glucanase

Fsβ-Glucanase mutants[a]	K_m (mg/ml) lichenan	k_{cat} (S^{-1})	k_{cat}/K_m (%)	Results and interpretation
Wild-type	1.9	871	100[b]	
Truncated	2.8	3911	269	Removal of the C-terminal domain and the enzyme is more stable and active
E56A[47]	ND	0	0	The carboxyl side-chain of E56 functions as the catalytic nucleophile. A56 and Q56 have no carboxyl group and no activity
E56D[47]	5.7	3.6	<1	
E56Q[47]	ND	0	0	
D58A[47]	ND	0	0	The carboxyl side-chain of D58 makes a H-bond to E56 to fix its orientation
D58E[47]	1.6	1.0	<1	The side-chain of E58 and N58 may partially retain the local structure but that of A58 cannot
D58N[47]	2.4	1.5	<1	
E60A[47]	ND	0	0	Functions as the general acid/base. The carboxyl side-chain of D60 retains residual activity but A60 and Q60 cannot
E60D[47]	2.6	1.6	<1	
E60Q[47]	ND	0	0	
M39F[47]	10	342	7	Located in the central hydrophobic core, involved in structural folding
W54F[48]	1.8	18	2	Make a H-bond to E56 to fix its orientation
W54Y[48]	1.3	23	3	
G63A[47]	3.1	315	22	Important for folding
W105F[48]	1.9	1018	103	Located between β-sheets. Not involved in substrate binding
W105H[48]	3.1	953	59	
W112F[48]	2.3	586	48	Not involved in substrate binding
W141F[48]	11.5	263	4	H-bond to E60 *via* a water molecule (W1 in Figure 6(B))
W141H[48]	12.8	43	<1	
W148F[48]	3.6	302	16	H-bond to Q70 and Q70 is further bridging to E60 *via* a water molecule (W2 in Figure 6(B))
W165F[48]	2.43	1125	89	Located in the convex site of the molecule
W165H[48]	1.8	987	105	Not involved in substrate binding or folding
W186F[48]	2.7	1846	130	Not involved in substrate binding
W198F[48]	3.6	1127	60	Not involved in substrate binding
W203F[48]	4.4	5476	242	Located in the substrate-binding pocket, likely involved in substrate binding
W203R[48]	17	44	<1	

[a] The mutant construction and activity measurement were described elsewhere.[47,48]

[b] The k_{cat}/K_m values for the mutants are expressed relative to the wild-type 1,3-1,4-β-D-glucanase (set at 100%) which has an absolute value of 518 ml^{-1} mg^{-1}. Reactions were performed with lichenan (8 mg/ml) as the substrate in 50 mM sodium citrate buffer (pH 6.0) or 50 mM phosphate buffer (pH 7.0) at optimum temperatures of the individual enzymes.

fine-tunes the structural fold. Figure 3 shows the superimposition of two β-glucanase structures from *F. succinogenes* and *B. licheniformis*. The average rms difference for the C^{α} atoms of 214 residues between the two structures, after least-squares fit, is 3.6 Å. The largest variations are located in the interconnecting loops and N and C-terminal regions. With these loops theoretically removed, the average rms difference between the two proteins for the 124 core residues located in β-sheet regions is only 1.3 Å.

Calcium ion binding site

A calcium ion is located on the edge of the convex side of the protein molecule, bound with nearly perfect pentagonal-bipyrimidal geometry to three backbone carbonyl oxygen atoms (Asn164, Asn189 and Gly222), an amide side-chain oxygen ($O^{\delta 1}$) atom (Asn164) and three water molecules (Figure 4(A)). The carbonyl oxygen atom of Gly222 and a water molecule are bound at the apex positions, while the other ones are arranged in the pentagonal plane. The calcium ion is located on the opposite surface of the active site, distant from the catalytic residue Glu56 (~25 Å). A calcium or cadmium ion was also identified at similar locations in several glycoside hydrolases of clan GH-B family 16, including *Bacillus* β-glucanases and κ-carrageenase.[55]

In *Bacillus* structures, the Ca^{2+} is located in a similar region but contains two different observed coordinations: (1) a pentahedral-bipyramidal geometry[41] coordinated to three backbone carbonyl oxygen atoms (Pro9, Gly45, and Asn207), an amide side-chain oxygen ($O^{\delta 1}$) atom (Asn207) and three water molecules; and (2) an octahedral geometry with the same residues but only two water molecules. The side-chain from the asparagine residues aligned at the same position, Asn164 in TFsβ-glucanase and Asn207 in *B. licheniformis*, is used in both structures to bind calcium ions. Also, the carbonyl backbone atoms from the residues aligned at the same positions, Asn164/Asn207,

Asn189/Pro9 and Gly222/Gly45, participate in calcium binding. Therefore, we expect that the calcium ion in TFsβ-glucanase may function in a similar manner to that in *Bacillus* enzymes.

The calcium binding to the engineered *Bacillus* glucanase (H(A16-M)) has been shown previously to stabilize the three-dimensional structure of the protein by guanidinium chloride unfolding and thermal inactivation experiments.[59] The melting point of TFsβ-glucanase in the presence and absence of calcium ions was measured using circular dichroism (CD) spectroscopy and it was found that the addition of 1 mM Ca^{2+} increased the melting point from 47.5–49.5°C (Figure 4(B)). The effect of calcium ions on thermal stability was further investigated by pre-incubating TFsβ-glucanase at various temperatures in the absence or presence of calcium ions. After heat treatment the residual enzymatic activities were higher for those of glucanases heated in the presence of calcium ions (Figure 4(C)). TFsβ-glucanase still retained 77% activity after the enzyme had been co-incubated with 50 mM $CaCl_2$ at 50 °C for ten minutes, whereas only 50% activity was observed for the enzyme pre-incubated in buffer solutions containing no Ca^{2+}. Therefore, these results suggest that calcium ions indeed slightly increase the thermal stability of the enzyme.

Substrate binding site and catalytic active site

The active site in TFsβ-glucanase is located in a cleft on the concave site of the β-jellyroll structure (Figure 1). The substrate-binding cleft has a length of ~28 Å and a width of ~8 Å, capable of accommodating approximately five glucopyranose units. Figure 5 shows the electrostatic potential map superimposed onto the molecular surface of TFsβ-glucanase. The bottom floor of the cleft is more acidic, containing several acidic residues, including Glu56, Asp58 and Glu60, which play catalytic roles in hydrolysis.[47] However, the wall of the cleft is more hydrophobic, lined with a number of aromatic residues, which may bind the hydrophobic faces of pyranoside rings as has been observed in many carbohydrate binding proteins.[60] A penta-saccharide molecule was modeled into the substrate-binding cleft of TFsβ-glucanase as shown in Figure 5. It is likely that residues Glu11, Tyr42, Asn44, Glu47, Trp54, Glu56, Asp58, Glu60, Gln70, Gln72, Glu85, Arg137, Asn139, Trp141, Trp148 and Trp203 are involved in substrate binding.

In the crystal structure of *Bacillus* glucanase,[42] an irreversible inhibitor 3,4-epoxybutyl β-D-cellobioside is covalently bound to the side-chain of Glu105, which is aligned with Glu56 in TFsβ-glucanase. This glutamate residue is conserved amongst all the 1,3-1,4-β-D-glucanases and likely functions as a catalytic nucleophile.[47] Glu56 in TFsβ-glucanase is located at the β4, forming hydrogen bonds with the side-chain of Trp54 and Asp58 (see Figure 2). The side-chain of Glu60 probably functions as the general acid/base[47] in the hydrolysis reaction and makes a hydrogen bond to the side-chain of Gln70 *via* a water molecule (see Figure 6). Therefore these catalytic residues are well positioned and their side-chain conformations are fixed by a hydrogen bond network. The distance between the carboxyl groups of Glu56 and Glu60 is 5.7 Å, which supports a mechanism of two SN2 reactions during hydrolysis resulting in the retention of protein configuration.[37] The residues namely Val55, Val57 and Ile59, located in between the catalytic residues are all hydrophobic and oriented away from the concave side, pointing toward the interior of the protein core between β-sheets. Although the overall protein backbone of TFsβ-glucanase has an rmsd of 3.6 Å from that of the *B. licheniformis* glucanase, the key residues involved in catalysis and substrate binding superimpose excellently (see Figure 6(A)), especially the three residues Glu56, Asp58 and Glu60, whose terminal oxygen atoms fit with the corresponding atoms with an average rms difference of only 0.45 Å. This result, in a good agreement with previous reports,[34] demonstrates that enzymes bearing similar catalysis mechanisms from different species usually have more conservative structural architecture in their active sites, but the overall protein structures are more deviated.

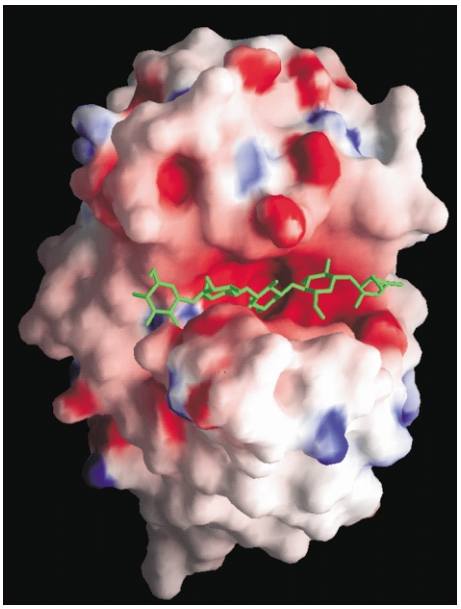

Figure 5. The electrostatic surface potential of TFsβ-glucanase was mapped onto the molecular surface and contoured at ±15 kT/e using GRASP.[68] Blue patches represent positive potential, red represents negative potential and white surface is neutral. A penta-saccharide molecule, colored in green, was modeled as the substrate bound at the active-site channel of TFsβ-glucanase. The substrate was constructed based on the structure of the inhibitor, 3,4-epoxybutyl β-D-cellobioside,[42] bound in a *Bacillus* glucanase and the model was briefly energy-minimized using CNS.

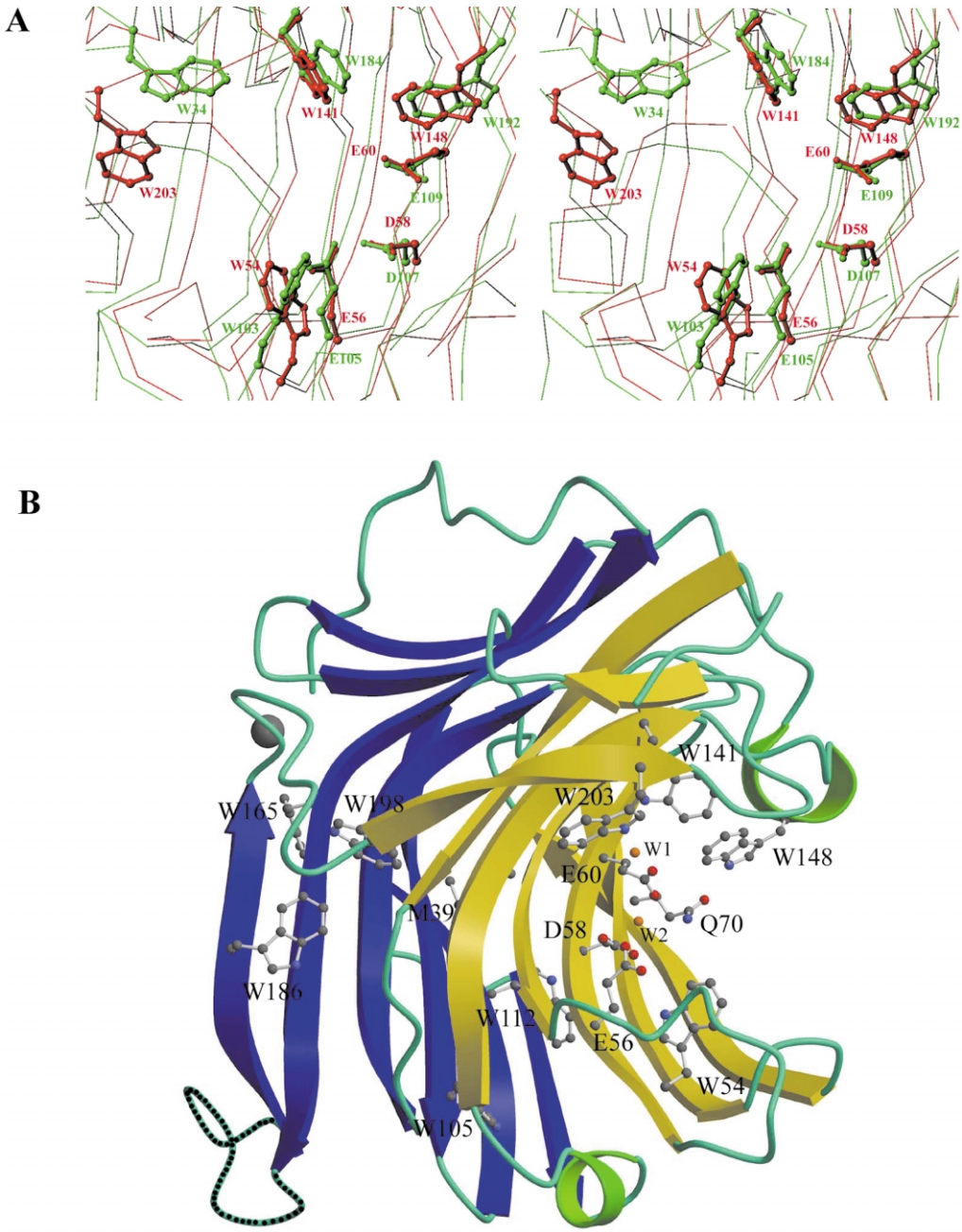

Figure 6. (A) The superimposition of the active site residues in TFsβ-glucanase (red) and β-glucanases from *B. licheniformis* (PDB ID:1gbg) (green). (B) The overall structure of TFsβ-glucanase with the residues listed in Table 2 for mutational study displayed in ball-and stick models. Water molecules involved in hydrogen bond networks are displayed as orange spheres. The water molecule W1 bridges Glu60 and Trp141 and W2 bridges Glu60 and Gln70.

Functional implications

The truncated form of Fsβ-glucanase shows a 2.6-fold increase in overall catalytic activity for lichenan compared to the wild-type enzyme (see Table 2). The C-terminal domain of Fsβ-glucanase (residues 231–339), containing a peptide sequence of P–X–S–S–S–S repeated five times, is not homologous to any other bacterial glucanase. The function of this C-terminal domain is intriguing and not known. From the crystal structure of the truncated Fsβ-glucanase, which shows a globular and compact fold, the C-terminal region is likely folded into a separate domain located at the convex side of the enzyme. Thus the C-terminal domain of Fsβ-glucanase is not directly involved in substrate binding and catalysis but it may regulate the activity of the enzyme through indirect effects, such as changing enzyme stability or rigidity.

A series of Fsβ-glucanase mutants have been constructed previously using site-directed

mutagenesis.[47,48] The functional implications for the phenotypes of the mutants based on this structural work are summarized in Table 2. In the crystal structure of Fsβ-glucanase, the three acidic residues located in the substrate-binding cleft, Glu56, Asp58 and Glu60, are involved in a hydrogen-bond network (see Figure 6). The carboxyl side-chain of Glu56 forms hydrogen bonds with Asp58 ($O^{\delta 1}$) and Trp54 ($N^{\epsilon 1}$). The carboxyl side-chain of Glu60 is bridged to Trp148 ($N^{\epsilon 1}$) and to Gln70 ($O^{\epsilon 1}$) *via* water molecules. In *Bacillus* β-glucanases, the residue equivalent to Glu56 has been proposed to function as a catalytic nucleophile and that equivalent to Glu60 to function as a general acid/base.[40] The mutation of Glu56 to aspartate retains the carboxyl side-chain for nucleophilic attack, therefore, the mutant E56D still retains residual activity, but the mutations of Glu56 to alanine (E56A) or glutamine (E56Q) completely abolish enzymatic activity. Similarly the isofunctional mutation of Glu60 to aspartate shows residual activity, but the mutants of E60A and E60Q without the carboxyl side-chain have no detectable enzyme activity. However, the carboxyl side-chain of Asp58 is not of crucial importance, in that of the mutations to it, only D58A has no detectable activity, whereas D58E and D58N contain residual activity probably because D58E and D58N are both capable of retaining the hydrogen bond with Glu56 using the carboxyl and amide side-chains, respectively. Therefore both the crystal structure and mutational results of Fsβ-glucanase support the earlier proposed mechanism that Glu56 and Glu60 are the two most critical catalytic residues involved in nucleophilic attack and proton transfer, and that Glu58 plays a structural role in stabilizing the catalytic residue and an electrostatic role in affecting the pK_a of the nucleophile residue Glu56.[40,43,61]

Several tryptophan and hydrophobic residues of Fsβ-glucanase have been mutated, and their enzymatic activities are listed in Table 2. Of these mutants, Trp165, Trp186 and Trp198 are located on the convex side of the enzyme and not involved in structural folding or substrate binding. Therefore these mutants, W165F, W165H, W186F and W198F only show minor changes in activity as compared to the wild-type enzyme. Mutations to several residues buried in the hydrophobic core of the enzyme, including Met39, Trp105 and Trp112 result in varied phenotypes. M39F has the lowest activity as compared to those of W105F, W105H and W112F, consistent with structural analysis that reveals that Met39 is located between the two β-sheets directly involved in structural folding. The tryptophan mutants with lower activity are all located on the concave side of the enzyme, including Trp54, Trp141, Trp148 and Trp203. These results are consistent with our crystal structure that suggests these tryptophan residues are probably involved in substrate binding, since they are located at the substrate binding cleft and most of them are also involved in stabilizing active site residues. Trp54 makes a weak hydrogen bond to the nucleophilic residue Glu56, Trp141 makes a hydrogen bond to the general acid residue Glu60 *via* a water molecule, Trp148 makes a hydrogen bond to Gln70 and Gln70 is further bridged to Glu60 *via* a water molecule. Moreover, the replacement of Trp203, with a phenylalanine or arginine, produces the opposite effect in enzyme activity, indicating that this residue may participate in hydrolysis and that a hydrophobic side-chain is preferred at this position. Replacement of Trp203 with a less bulky side-chain residue, phenylalanine, effectively facilitated the catalytic efficiency of the enzyme, whereas, introducing a positive charge group in residue 203 apparently interrupts the catalytic function of the enzyme. In the crystal structure, Trp203 is located at the one end of the cleft, very likely involved in substrate binding.

Conclusions

The crystal structure of the truncated Fsβ-glucanase from *F. succinogenes* represents the first structure of a natural circularly permuted enzyme from glycoside hydrolase family 16. The crystal structure of TFsβ-glucanase has a β-jellyroll topology similar to those of *Bacilli* enzymes. However, the structures of the *de novo* engineered circularly permuted *Bacilli* glucanases are more similar to their highly homologous parent enzymes than to TFsβ-glucanase, with analogous starting and ending amino acid termini. This result suggests that the protein structure relies more on sequence identity than topology. Taken together with the mutational results, the crystal structure of TFsβ-glucanase provides important information that allows for a better understanding of the structure and functional relationship of this enzyme and may be useful, in the future, for designing enzymes with improved catalytic efficiency or stability for industrial application.

Materials and Methods

Expression and purification of TFsβ-glucanase

The DNA coding sequence of the truncated form of Fsβ-glucanase was obtained by using a PCR-based method with a pair of specific primers and the full-length cDNA of Fsβ-glucanase as template.[47] The amplified DNA fragment, which encoded the N-terminal domain from residues 1 to 258 of full length Fsβ-glucanase, was ligated with a pET26b(+) vector (Novagen, USA) and then transformed into B834(DE3) host cells. The cells were cultured in minimal M10 medium supplemented with glucose, (8 g/l), 19 L-amino acids mixture (80 mg/l), seleno-L-methionine (80 mg/l) and kanamycin (30 mg/l). The truncated form of Fsβ-glucanase was effectively expressed and secreted into M10 culture medium as a soluble protein at 33 °C after 1 mM IPTG induction for 16 hours. The standard defined medium M10, contained the chemicals in g/l as follows:

NH_4Cl, 2; KH_2PO_4, 6; Na_2HPO_4, 12; $MgSO_4.7H_2O$, 0.5; $FeSO_4.7H_2O$, 0.025; and 1 ug/ml of riboflavin, niacin amide, pyridoxine monohydrochloride and thiamine. The culture supernatant containing approximately 85% of the expressed truncated form of Fsβ-glucanase was collected by centrifugation at 8000*g* for 15 minutes at 4 °C and concentrated to 0.1 volume using a Pellicon Cassette concentrator (Millipore, Bedford, MA) with a 10,000 M_r cut-off membrane. The concentrated supernatant was then dialyzed against buffer A (50 mM Tris–HCl buffer pH 7.5) and loaded onto a Q-Sepharose FF (Amersham Pharmacia Biotech) column pre-equilibrated with the same buffer. The truncated form of Fsβ-glucanase was collected from eluants of the column using a 0–1 M NaCl salt gradient in buffer A. Second and third Ni-NTA affinity columns, equilibrated with buffer B (50 mM sodium phosphate (pH 8.0), 0.3 M NaCl, 10 mM imidazole) were then employed for further purification of the truncated enzymes. From a 10–300 mM imidazole gradient eluant, a homogeneous enzyme preparation was obtained, as verified by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (data not shown). TFsβ-glucanase has a molecular mass of 28,754 Da, as determined by mass spectrometry, corresponding to 258 amino acid residues (calculated molecular mass: 28,750 Da).

Measurement of melting point

The thermal denaturations of TFsβ-glucanase were monitored by CD on a Jasco J715 CD spectrometer at temperatures ranging from 25 °C to 80 °C. Spectra were collected from 200 nm to 260 nm in 1.3 nm increments and the melting point was derived from the CD signals recorded at 214 nm. The protein concentration was adjusted to 0.5 mg/ml in buffer solutions of 50 mM sodium phosphate at pH 7.0 in the absence or presence of 50 mM $CaCl_2$. The fraction of folded protein (F_{app}) as monitored by $CD_{214\ nm}$ was calculated by $F_{app} = (Y_{obsd} - Y_U)/(Y_N - Y_U)$.[62] Y_{obsd} represents the observed value of CD at 214 nm of TFsβ-glucanase at various temperatures. Y_N and Y_U represent the CD values at 214 nm of the TFsβ-glucanase at 25 °C and 70 °C, respectively. The melting point was the temperature where 50% of the protein was unfolded ($F_{app} = 0.5$).

Effect of metal ions on the activity and thermostability

The effect of calcium ions on the thermostability of TFsβ-glucanase was studied by treating the enzyme at 40, 45, 50, 55, 60, 70 and 80 °C, respectively, for ten minutes, in the presence (1 mM or 50 mM $CaCl_2$) or absence of the metal ions. The residual enzyme activities after the treatment were then measured by determining the rate of reducing sugar production from the hydrolysis of the substrate (lichenan) on the basis of the method described elsewhere.[47]

Crystallization and data collection

Crystallization of the seleno-methionine-labeled TFsβ-glucanase was carried out using the hanging-drop vapor-diffusion method at room temperature. Prior to crystallization, the purified protein was concentrated to 10 mg/ml in 10 mM Tris–HCl buffer (pH 7.5). Drops of 1 μl of the protein solution were mixed with 1 μl of reservoir solution containing 2 mM $CaCl_2$, 0.1 M CH_3COONa, 0.05 M Tris–HCl (pH 9.0) and 30% (w/v) PEG4000. Crystals appeared after three days and reached a final size of about 0.1 mm × 0.2 mm × 0.5 mm within a week.

Crystals were soaked in cryo-protectant, consisting of 10% (v/v) glycerol in reservoir solution, for one minute before data collection. MAD data sets were collected at a low temperature (~ − 150 °C) using the synchrotron X-ray radiation source at beam line BL18B equipped with a Quantum-4 CCD detector in Photon Factory (Tuskuba, Japan). Diffraction data were recorded from the crystal at: (a) point of inflection (λ1 = 0.9795 Å); (b) peak of absorption (λ2 = 0.9793 Å); and (c) remote (λ3 = 0.9600 Å). High quality X-ray diffracted data sets were obtained from a single crystal up to a resolution of 1.7 Å (Table 1). Indexing and integration of diffraction data were performed using DENZO and SCALEPACK.[63] The space group was determined to be orthorhombic $P2_12_12_1$ with cell dimensions of $a = 40.87$ Å, $b = 73.27$ Å and $c = 73.75$ Å, containing one molecule per asymmetric unit.

Structure determination, model building and refinement

The crystal structure of TFsβ-glucanase was solved by the MAD method using the anomalous signals from a seleno-methionine-labeled protein. The five expected selenium sites in residues 1, 27, 29, 39 and 223 were identified using programs CNS[64] and SOLVE.[65] Four selenium sites (residues 27, 29, 39 and 223) were located using an automated Patterson heavy-atom search in CNS, whereas four selenium sites (residues 1, 27, 39 and 223) were located by SOLVE. Positions and occupancies of the four selenium atoms identified in CNS were further refined. These four sites had good occupancy (1.0–1.3) and reasonable temperature factors (~10 Å^2) and gave good phasing statistics, as listed in Table 1. At this point, the initial phases were improved by density modification techniques in CNS and the first Fourier map was calculated (see Figure 2), clear enough for tracing most of the polypeptide chain, except residues 41–45 and 171–181. Subsequent structure refinement with CNS involved careful model building (using TURBO-FRODO[66]), simulated annealing, positional and *B*-factor refinement and the addition of water molecules that were carried out using the remote data set. The final structural model contained residues 1–243, 280 water molecules and a calcium ion with a final *R*-factor of 19.2% and *R*-free of 23.8% for all the reflections ($F > 0$) in the resolution range of 40–1.7 Å (Table 1).

Protein Data Bank accession codes

The coordinates and structural factors have been deposited in the RCSB Protein Data Bank under accession code 1mve.

Acknowledgements

This work was supported by research grants from Academia Sinica and the National Science Council of the Republic of China to H. S. Y. and L.-F. S.

References

1. Uliel, S., Fliess, A. & Unger, R. (2001). Naturally occurring circular permutations in proteins. *Protein Eng.* **14**, 533–542.
2. Cunningham, B. A., Hemperley, J. J., Hopp, T. P. & Edelman, G. M. (1979). Favin versus concanavalin A: circularly-permutated amino acid sequences. *Proc. Natl Acad. Sci. USA*, **76**, 3218–3222.
3. Lindqvist, Y. & Schneider, G. (1997). Circular permutations of natural protein sequences: structural evidence. *Curr. Opin. Struct. Biol.* **7**, 422–427.
4. Hahn, M., Piotukh, K., Borriss, R. & Heinemann, U. (1994). Native-like *in vivo* folding of a circularly permutated jellyroll protein shown by crystal structure analysis. *Proc. Natl Acad. Sci. USA*, **91**, 10417–10421.
5. Murzin, A. G. (1998). Probable circular permutation in the flacin-binding domain. *Nature Struct. Biol.* **5**, 101.
6. Jeltsch, A. (1999). Circular permutations in the molecular evolution of DNA methyltransferases. *J. Mol. Evol.* **49**, 161–164.
7. Rojas, A. & Romeu, A. (1996). A sequence analysis of the β-glucosidase sub-family B. *FEBS Letters*, **378**, 93–97.
8. Xu, S.-Y., Xiao, J.-P., Posfai, J., Maunus, R. & Benner, J. (1997). Cloning of the BssHII restriction-modification system in *Escherichia coli*: BssHII methyltransferase contains circularly permuted sytosine-5 methyltransferase motifs. *Nucl. Acids Res.* **25**, 3991–3994.
9. Pointing, C. P. & Russell, R. B. (1995). Swaposins: circular permutations within genes encoding saposin homologues. *Trends Biochem. Sci.* **20**, 179–180.
10. MacGregor, E. A., Jespersen, H. M. & Svensson, B. (1996). A circularly permuted α-amylase-type α/β-barrel structure in glucan-synthesizing glucosyltransferases. *FEBS Letters*, **378**, 263–266.
11. Jia, J., Huang, W., Schorken, U., Sahm, H., Sprenger, G. A., Lindqvist, Y. & Schneider, G. (1996). Crystal structure of transaldolase B from *Escherichia coli* suggests a circular permutation of the α/β barrel within the class I aldolase family. *Structure*, **4**, 715–724.
12. Castillo, R. M., Mizuguchi, K., Dhanaraj, V., Albert, A., Blundell, T. L. & Murzin, A. G. (1999). A six-stranded double-φ β-barrel is shared by several protein superfamilies. *Struct. Fold. Des.* **7**, 227–236.
13. Polekhina, G., Board, P. G., Gali, R. R., Rossjohn, J. & Parker, M. W. (1999). Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. *EMBO J.* **18**, 3204–3213.
14. Jung, J. & Lee, B. (2000). Protein structure alignment using environmental profiles. *Protein Eng.* **13**, 535–543.
15. Antcheva, N., Pintar, A., Patthy, A., Simoncsits, A., Barta, E., Tchorbanov, B. & Pongor, S. (2001). Proteins of circularly permuted sequence present within the same organism: the major serine proteinase inhibitor from *Capsicum annuum* seeds. *Protein Sci.* **10**, 2280–2290.
16. Jung, J. & Lee, B. (2001). Circularly permutated proteins in the protein structure database. *Protein Sci.* **10**, 1881–1886.
17. Einspahr, H., Parks, E. H., Suguna, K. & Subramanian, E. (1986). The crystal structure of pea lectin at 3.0 Å resolution. *J. Biol. Chem.* **261**, 16518–16527.
18. Edelman, G. M., Cunningham, B. A., Reekem, G. N., Jr, Becker, J. W., Waxdal, M. J. & Wang, J. L. (1972). The covalent and three-dimensional structure of concanavalin A. *Proc. Natl Acad. Sci. USA*, **69**, 2580–2584.
19. Hardman, K. D. & Ainsworth, C. F. (1972). Structure of concanavalin A at 2.4 Å resolution. *Biochemistry*, **11**, 4910–4919.
20. Bowles, D. J., Marcus, S. E., Pappin, D. J. C., Findlay, J. B. C., Eliopoulus, E., Maycox, P. R. & Burgess, J. (1986). Posttranslational processing of concanavalin A precursors in Jackbean cotyledons. *J. Cell Biol.* **102**, 1284–1297.
21. Carrington, D. M. & Auffret, A. E. H. D. (1985). Polypepeptide ligation occurs during post-translational modification of concanavalin A. *Nature*, **313**, 64–67.
22. Heinemann, U. & Hahn, M. (1995). Circular permutations of protein sequence: not so rare? *Trends Biochem. Sci.*, 349–350.
23. Anderson, M. A. & Stone, B. A. (1975). A new substrate for investigating the specificity of β-glucan hydrolases. *FEBS Letters*, **52**, 202–207.
24. Borriss, R., Manteuffel, R. & Hofemeister, J. (1988). Molecular cloning of a gene coding for thermostable betaglucanase from *Bacillus macerans*. *J. Basic Microbiol.* **28**, 3–10.
25. Bueno, A., Vazquez de ALdana, C. R., Correa, J., Villa, T. G. & del Rey, F. (1990). Synthesis and secretion of a Bacillus circulans WL-12 1,3-1,4-β-D-glucanase in *Escherichia coli*. *J. Bacteriol.* **172**, 2160–2167.
26. Gosalbes, M. J., Perez-Gonzalez, J. A., Gonzalez, R. & Navarro, A. (1991). Two β-glucanase genes are clustered in *Bacillus polymyxa*: molecular cloning, expression,and sequence analysis of genes encoding a xylanase and an *endo*-β-(1,3)-(1,4)-glucanase. *J. Bacteriol.* **173**, 7705–7710.
27. Hofemeister, J., Kurtz, J., Borriss, R. & Knowles, J. (1986). The β-glucanase gene from *Bacillus amyloliquefaciens* shows extensive homology with that of *Bacillus subtilis*. *Gene (Amst.)*, **49**, 177–187.
28. Lloberas, J., Perez-Pons, J. A. & Querol, E. (1991). Molecular cloning, expression and nucleotide sequence of the *endo*-β-1,3-1,4 -D-glucanase-encoding gene from *Bacillus licheniformis*. *Eur. J. Biochem.* **197**, 337–343.
29. Louw, M. E., Reid, S. J. & Watson, T. G. (1993). Characterization, cloning, expression and sequencing of a thermostable endo-(1,3-1,4) β-glucanase-encoding gene from alkalophilic *Bacillus brevis*. *Appl. Microbiol. Biotechnol.* **38**, 507–513.
30. Murphy, N., McConnell, D. J. M. & Cantwell, B. A. (1984). The DNA sequence of the gene and genetic control sites for the excreted *B. subtilis* enzyme β-glucan hydrolase. *Nucl. Acids Res.* **12**, 5355–5367.
31. Tzuka, H., Yuuki, T. & Yabuuchi, S. (1989). Construction of a β-glucanase hyperproducing *Bacillus subtilis* using the cloned β-glucanase gene and a multi-copy plasmid. *Agric. Biol. Chem.* **53**, 2335–2339.
32. Teather, R. M. & Erfle, J. D. (1990). DNA sequence of a *Fibrobacter succinogenes* mixed-linkage β-glucanase (1,3-1,4)-β-D-glucan 4-glucanohydrolase gene. *J. Bacteriol.* **172**, 3837–3841.
33. Flint, H. J., Martin, J., McPherson, C. A., Daniel, A. S. & Zhang, J.-X. (1993). A bifunctional enzyme with seperate xylanase and β(1,3-1,4)-glucandomains encoded by the xynD gene of *Ruminococcus flavefaciens*. *J. Bacteriol.* **175**, 2943–2951.
34. Schimming, S., Schwarz, W. H. & Staudenbauer, W. L. (1992). Structure of the *Clostridium thermocellum* gene licB and the encoded β-1,3-1,4-glucanase: a catalytic

region homologous to *Bacillus lichenases* joined to the reiterated domain of clostridial cellulases. *Eur. J. Biochem.* **204**, 13–19.
35. Fincher, G. B., Lock, P. A., Morgan, M. M., Lingelbach, K., Wettenhall, R. E. H., Mercer, J. F. K. *et al.* (1986). Primary structure of the (1,3-1,4)-β-D-glucan 4-glucanohydrolase from barley aleurone. *Proc. Natl Acad. Sci. USA.* **83**, 2081–2085.
36. Litts, J. C., Simmons, C. R., Karrer, E. E., Huang, N. & Rodriguez, R. L. (1990). The isolation and characterization of a barley 1,3-1,4-β-glucanase gene. *Eur. J. Biochem.* **194**, 831–838.
37. Henrissat, B. & Davies, G. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, **3**, 853–859.
38. Malet, C., Jimenez-Barvero, J., Bernabe, M., Brosa, M. & Planas, A. (1993). Stereochemical course and structure of the products of the enzymatic action of endo-1,3,-1,4-β-D-glucan 4-glucanohydrolase from *Bacillus licheniformis*. *Biochem. J.* **296**, 753–758.
39. Henrissat, B. & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**, 695–696.
40. Hahn, M., Olsen, O., Politz, O., Borriss, R. & Heinemann, U. (1995). Crystal structure and site-directed mutagenesis of *Bacillus macerans endo*-1,3-1,4-β-glucanase. *J. Biol. Chem.* **270**, 3081–3088.
41. Hahn, M., Pons, J., Planas, A., Querol, E. & Heinemann, U. (1995). Crystal structure of *Bacillus licheniformis* 1,3-1,4-β-D-glucan 4-glucanohydrolase at 1.8 Å resoltuion. *FEBS Letters*, **374**, 221–224.
42. Keitel, T., Simon, O., Borriss, R. & Heinemann, U. (1993). Molecular and active-site structure of a *Bacillus* 1,3-1,4-β-glucanase. *Proc. Natl Acad. Sci. USA*, **90**, 5287–5291.
43. Hahn, M., Keitel, T. & Heinemann, U. (1995). Crystal and molecular structure at 0.16 nm resolution of the hybrid *Baccillus endo*-1,3-1,4-β-D-glucan 4-glucanohydrolase H(A16-M). *Eur. J. Biochem.* **232**, 849–858.
44. Ay, J., Hahn, M., Decanniere, K., Piotukh, K., Borriss, R. & Heinemann, U. (1998). Crystal structures and properties of de novo circularly permutated 1,3-1,4-β-glucanase. *Proteins: Struct. Funct. Genet.* **30**, 155–167.
45. Viguera, A. R., Serrano, L. & Wilmanns, M. (1996). Different folding transition states may result in the same native structure. *Nature Struct. Biol.* **3**, 874–880.
46. Tsai, L.-C., Shyur, L.-F., Lin, S.-S. & Yuan, H. S. (2001). Crystallization and preliminary X-ray diffraction analysis of the 1,3-1,4-β-D-glucanase from *Fibrobacter succinogenes*. *Acta Crystallog.* D**57**, 1303–1306.
47. Chen, J.-L., Tsai, L.-C., Wen, T.-N., Tang, J.-B., Yuan, H. S. & Shyur, L.-F. (2001). Directed mutagenesis of specific active site residues on *Fibrobacter succinogenes* 1,3-1,4-β-D-glucanase significantly affects catalysis and enzyme structural stability. *J. Biol. Chem.* **276**, 17895–17901.
48. Cheng, H.-L., Tsai, L.-C., Lin, S.-S., Yuan, H. S., Yang, N.-S., Lee, S.-H. & Shyur, L.-F. (2002). Mutagenesis of Trp54 and Trp203 residues on *Fibrobacter succinogenes* 1,3-1,4-β-D-glucanase significantly affects catalytic activities of the enzyme. *Biochemistry*, **41**, 8759–8766.
49. Laskowski, R. A., MacArthur, M. V., Moss, D. S. & Thorton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
50. Divne, C., Stahlberg, J., Reinikainen, T., Rhohomen, L., Pattersson, G., Knowles, J. K. C. *et al.* (1994). The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science*, **265**, 524–528.
51. Torronen, A., Harkki, A. & Rouvinen, J. (1994). Three-dimensional structure of endo-1,4-beta-xylanase II from *Trichoderma reesei*: two conformational states in the active site. *EMBO J.* **13**, 2493–2501.
52. Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J. & Yaguchi, M. (1994). Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* **3**, 467–475.
53. Crennell, S. J., Hreggvidsson, G. O. & Karlsson, E. N. (2002). The structure of *Rhodothermus marinus* Cel12A, a highly thermostable family 12 endoglucanase, at 1.8 Å resolution. *J. Mol. Biol.* **320**, 883–897.
54. Sulzenbacher, G., Mackenzie, L. F., Wilson, K. S., Withers, S. G., Dupont, C. & Davies, G. J. (1999). The crystal structure of a 2-fluorocellotriosyl complex of the *Streptomyces lividans* endoglucanase CelB2 at 1.2 Å resolution. *Biochemistry*, **38**, 4826–4833.
55. Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B. & Dideberg, O. (2002). The κ-carrageenase of *P. carrageenovora* features a tunnel-shapped active site: a novel insight in the evolution of clan-B glycoside hydrolases. *Structure*, **9**, 513–525.
56. Szabo, L., Jamal, S., Xie, H., Charnock, S. J., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2001). Structure of a family 15 carbohydrate-binding module in complex with xylopentaose, evidence that xylan binds in an approximate 3-fold helical conformation. *J. Biol. Chem.* **276**, 49061–49065.
57. Johnson, P. E., Creagh, A. L., Brun, E., Joe, K., Tomme, P., Haynes, C. A. & McIntosh, L. P. (1998). Calcium binding by the N-terminal cellulose-binding domain from *Cellulomonas fimi* β-1,4-glucanase CenC. *Biochemistry*, **37**, 12772–12781.
58. Czjzek, M., Bolam, D. N., Mosbah, A., Allouch, J., Fontes, C. M. G. A., Ferreira, L. M. A. *et al.* (2001). The location of the ligand-binding site of carbohydrate-binding modules that have evolved from a common sequence is not conserved. *J. Biol. Chem.* **276**, 48580–48587.
59. Keitel, T., Meldgaard, M. & Heinemann, U. (1994). Cation binding to a *Bacillus* (1,3-1,4)-β-glucanase-Geometry, affinity and effect on protein stability. *Eur. J. Biochem.* **222**, 203–214.
60. Vyas, N. K. (1991). Atomic features of protein-carbohydrate interactions. *Curr. Opin. Struct. Biol.* **1**, 732–740.
61. Juncosa, M., Pons, J. D. T., Querol, E. & Planas, A. (1994). Identification of active site carboxylic residues in *Bacillus licheniformis* 1,3-1,4-β-D-glucan 4-glucanohydrolase by site-directed mutagenesis. *J. Biol. Chem.* **269**, 14530–14535.
62. Cai, K. & Schirch, V. (1996). Structural studies on folding intermediates of serine hydroxymethyltransferase using single tryptophan mutants. *J. Biol. Chem.* **271**, 2987–2994.
63. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
64. Grosse-Kunstleve, R. W. & Brunger, A. T. (1999). A highly automated heavy-atom search procedure for macromolecular structures. *Acta Crystallog.* **55**, 1568–1577.
65. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog.* D**55**, 849–861.

66. Roussel, A. & Cambillau, C. (1992). *The TURBO-FRODO graphics package* Silicon Graphics Geometry Partners Directory, Vol. **81**, Silicon Graphics, Mountain View, CA.
67. Barton, G. J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.
68. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.

Edited by R. Huber

(Received 14 January 2003; received in revised form 7 May 2003; accepted 7 May 2003)