Crystal structure of *Bacillus licheniformis* 1,3-1,4-β-D-glucan 4-glucanohydrolase at 1.8 Å resolution

Michael Hahn[a], Jaume Pons[b,c], Antoni Planas[b], Enrique Querol[c], Udo Heinemann[a,*]

[a]*Forschungsgruppe Kristallographie, Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle-Str. 10, 13122 Berlin, Germany*
[b]*CETS Institut Químic de Sarrià, Universitat Ramón Llull, 08017 Barcelona, Spain*
[c]*Institut de Biologia Fonamental, Universitat Autónoma de Barcelona, 08193 Bellatorra (Barcelona), Spain*

Received 28 August 1995; revised version received 20 September 1995

Abstract The crystal structure of the 1,3-1,4-β-D-glucan 4-glucanohydrolase from *Bacillus licheniformis* is solved at a resolution of 1.8 Å and refined to $R = 16.5\%$. The protein has a similar β-sandwich structure as the homologous enzyme from *Bacillus macerans* and the hybrid H(A16-M). This demonstrates that the jellyroll fold of these proteins is remarkably rigid and only weakly influenced by crystal contacts. The crystal structure permits to extend mechanistic considerations derived for the *B. licheniformis* enzyme to the entire class of bacterial 1,3-1,4-β-D-glucan 4-glucanohydrolases.

Key words: 1,3-1,4-β-D-Glucan 4-glucanohydrolase; 1,3-1,4-β-Glucanase; *Bacillus licheniformis*; Crystal structure; β-Glucan hydrolysis; Enzyme mechanism

1. Introduction

1,3-1,4-β-D-Glucan 4-glucanohydrolase (1,3-1,4-β-glucanases, E.C. 3.2.1.73) from different *Bacillus* strains have been cloned and sequenced [1–7]. These enzymes cleave β-1,4-glycosidic bonds that are adjacent to β-1,3-glycosidic bonds in mixed linked glucans. They show sequence similarities with β-1,3-glucanases (laminarinases) but not with 1,3-1,4-β-glucanases from plants [8]. The crystal structure of the 1,3-1,4-β-glucanase of *B. macerans* (BGLM) and the closely related hybrid *Bacillus* enzyme H(A16-M) are known as well as circularly permuted variants thereof [9–11]. 1,3-1,4-β-Glucanases belong to the jellyroll β-sandwich-type proteins, where the sheets of the sandwich are curved and create a channel for substrate binding and cleavage.

Several studies were undertaken to establish the mechanism of carbohydrate cleavage by 1,3-1,4-β-glucanases. In BGLL, the enzyme from *Bacillus licheniformis*, site-directed mutagenesis experiments indicated that Glu-105 acts as a nucleophile while Glu-109 is the general acid catalyst, and that Asp-107 is also crucial for activity[1] [12,13]. Likewise, crystallographic and site-directed mutagenesis experiments on BGLM [10] suggested that the catalytic amino acids lie on one β-strand and consist of a proton donor, Glu-107, equivalent to Glu-109 in BGLL, as well as a nucleophile, Glu-103, corresponding to Glu-105 in BGLL. The distance between both catalytic amino acids in the crystal structure of BGLM and H(A16-M) agrees with that expected for a retaining glycosidase (4.5–5.5 Å) as opposed to that for an inverting glycosidase (9–9.5 Å) [14]. Glycosidic bond hydrolysis has been suggested to take place by a double displacement reaction where the nucleophile stabilizes the development of a positive charge at the anomeric carbon in the transition state to yield either a covalent or oxo-carbonium intermediate before a water molecule completes the reaction which, in summary, results in net retention of the configuration of the anomeric carbon C_1.

That the reaction indeed proceeds with retention of configuration was shown biochemically for BGLL [15]. Since the three-dimensional structure of this protein was not known, it was unclear whether we were justified to assume identical enzyme mechanisms for BGLM and BGLL. Here we present the crystal structure of BGLL at high resolution and show its high degree of similarity with BGLM and H(A16-M). From this we conclude that the bacterial 1,3-1,4-β-glucanases share a very rigid jellyroll domain and that their enzymatic reactions are stereochemically identical, probably employing very similar reaction mechanisms.

2. Experimental

2.1. Crystallization and data acquisition

BGLL was lyophilized for storage and transport and redissolved in 10 mM HEPES pH 7.0 and 2 mM $CaCl_2$. It was concentrated to 44 mg/ml and mixed with an equal volume of precipitant solution in a hanging drop vapor diffusion setup. The precipitant solution contained 20% PEG 8000 and 2 mM $CaCl_2$ and was buffered with 50 mM potassium phosphate at pH 8.0. Crystals grew over night to dimensions of $0.6 \times 0.6 \times 0.2$ mm^3, and cell constants were determined to $a = 35.33$ Å, $b = 39.13$ Å, $c = 43.88$ Å, $\alpha = 64.66°$, $\beta = 105.86°$ and $\gamma = 110.68°$ in space group P1. There is one protein molecule per unit cell. Two diffraction data sets were measured with CuK_α radiation, one with a FAST area detector and one with a MAR-Research (Hamburg) Imaging Plate, both mounted on a FR-571 (Enraf-Nonius, Delft) rotating anode generator. They were processed with the MADNES [16] and MOSFLM

*Corresponding author. Fax: (49) (30) 9406 2548.

Abbreviations: BGLL, 1,3-1,4-β-glucanase from *Bacillus licheniformis*; BGLM, 1,3-1,4-β-glucanase from *Bacillus macerans*; BGLA, 1,3-1,4-β-glucanase from *Bacillus amylolique-faciens*; H(A16-M), hybrid enzyme with amino acids 1–16 derived from BGLA and amino acids 17–214 from BGLM; PEG, polyethylene glycol; rms, root mean square.

[1]The previously used amino acid numbering of BGLL includes the 29-amino acid sequence of the signal peptide, while that for BGLM was for the mature, processed enzyme. In this paper, the numbering of the mature enzyme is used for both BGLL and BGLM for easier comparison. Mature BGLL consists of 214 amino acid residues and has a relative molecular mass of 24418.

SSDI 0014-5793(95)01111-0

[17] software packages, respectively, and merged with the CONDENSE option of SHELX-76 [18].

2.2. Structure determination and refinement

The structure was solved by calculating the rotation function with X-PLOR [19]. The search model, H(A16-M) [20], gave a single peak 5 σ above the mean value. Rigid body refinement against diffraction data between 12 and 3 Å yielded an R value of 36.3%. The amino acids that differ in the sequence of H(A16-M) compared to BGLL were changed to alanine, and a simulated-annealing slowcool protocol was executed for all data between 10 and 1.8 Å. Afterwards, the sequence was corrected with O [21] on the graphics screen and the slowcool run was repeated, resulting in an R value of 29.2%. Further refinement was done with X-PLOR and TNT [22]. Solvent molecules were added to the coordinate set until all significant peaks disappeared from the difference electron density map. The O_γ atom of Ser-50 was refined in two positions with half occupancy each. The experimental data and atomic coordinates of the refined protein model were submitted to the Brookhaven Protein Databank (entry 1GBG).

3. Results and discussion

3.1. Protein structure

The final model of BGLL has 91% of its residues in the most favored regions of the Ramachandran diagram [23] and an R value of 16.5%. The free R value [24] after calculating a simulated annealing protocol with the final model is 23.6%. The geometrical parameters are of very good quality, as shown in Table 1. Secondary structural elements were identified with HERA [25] based on a DSSP [26] analysis. Only one small difference with respect to the H(A16-M) structure is revealed concerning the β-strand from residue 18 to 21, which extends from residue 17 to 21 in this protein. At position 211 of the enzyme a tyrosine residue was identified (Fig. 1) where a serine was expected from the published sequence [6].

The overall structure of BGLL is nearly identical to the known structures of the *Bacillus* 1,3-1,4-β-glucanases BGLM and H(A16-M) [10,11]. The fold is dominated by two antiparallel major β-sheets that consists of 8 and 7 strands, respectively (Fig. 2). As they are bent, they create a channel in which the substrate is cleaved. A third, minor sheet connects the major sheets showing a mixed parallel and antiparallel bonding scheme. The *cis*-proline at position 201 and the disulfide bond from Cys-32 to Cys-61, known from the model of H(A16-M), are conserved. This disulfide bond has a minor structural and functional role as shown in BGLL by reduction and site-directed replacement. Activity and stability are essential unaffected upon disulfide reduction, but protein stability drastically decreases in Cys-to-Ala mutants [27].

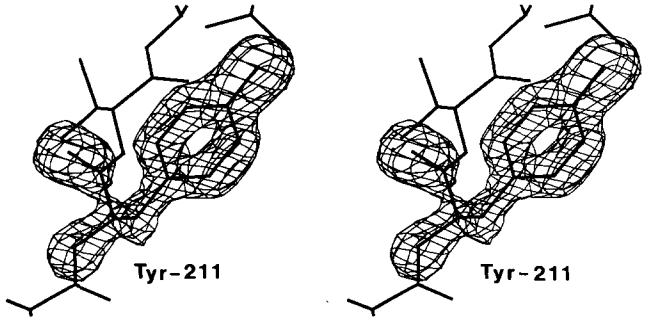

Fig. 1. Stereographic drawing of F_o–F_c difference density around residue Tyr-211 of BGLL. The omit map was calculated after deleting the tyrosine from the atom list. Drawn with SETOR [28].

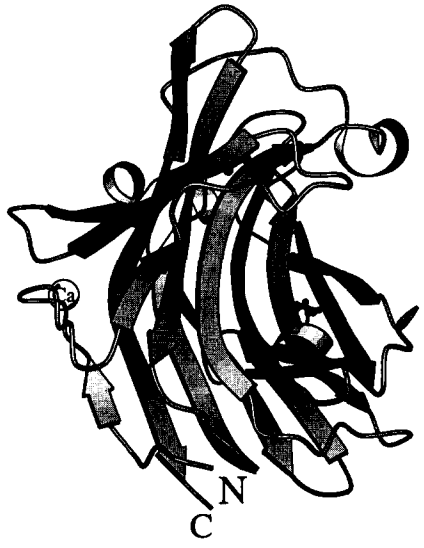

Fig. 2. Structure of BGLL with β-strands as arrows, α-helical turns as ribbons and coil regions as lines. The N- and C-termini as well as the calcium ion, the disulfide bond and the catalytic site residues Glu-105, Asp-107 and Glu-109 are shown. Drawn with MOLSCRIPT [29].

A Ca^{2+} ion is found in the same site as in H(A16-M), but it is coordinated in a pentahedral-bipyramidal manner. The carbonyl oxygen atoms of Pro-9 and Asn-207 and O_δ of Asn-207, together with two water molecules, form the pentahedral plane whereas the carbonyl oxygen of Gly-45 and another water molecule occupy the apical positions of the bipyramid. In H(A16- M) the coordination is octahedral, although a pentahedral-bipyramidal coordination is indicated in this model by a water molecule that was assigned to two mutually exclusive locations with half occupancy each.

Asn-28 lies in a disallowed region of the Ramachandran diagram. This can be explained by the crystal contact of this residue with amino acids 161 and 162 of a neighboring molecule. The carbonyl oxygen of Asn-28 is only 3.3 Å away from the carbonyl carbon of Asp-161 of the translationally related molecule, and its C_β approaches C_α of Gly-162 to within 3.4 Å. The conformation of Asn-28 is under these steric constraints, because of the strong hydrogen-bonding crystal contact formed in this region of the protein surface by the carbonyl oxygens of Asn-26 and Ser-25 with $N_{\eta 1}$ and $N_{\eta 2}$ of Arg-210. Besides these contacts there are hydrogen-bonding contacts to three other molecules in the crystal (Table 2).

Table 1
X-ray diffraction data and refinement results

Cell constants	a = 35.33 Å, b = 39.13 Å, c = 43.88 Å, α = 64.66°, β = 105.86°, γ = 110.68°
Space group	P1
Unique reflections	17,238
Completeness	94.6%
R_{merge} [a]	8.2%
Resolution range	6.0–1.8 Å
Rms bond deviation	0.004 Å
Rms angle deviation	1.8°
B factor correlation	3.2 Å^2
Solvent molecules	1 Ca^{2+} ion, 158 water molecules
R factor	16.5%

[a] $R_{merge} = \Sigma_{h,i} |I(h)_i - <I(h)>| / \Sigma_{h,i} I(h)_i$, where $I(h)_i$ are the measurements contributing to the mean reflection intensity, $<I(h)>$.

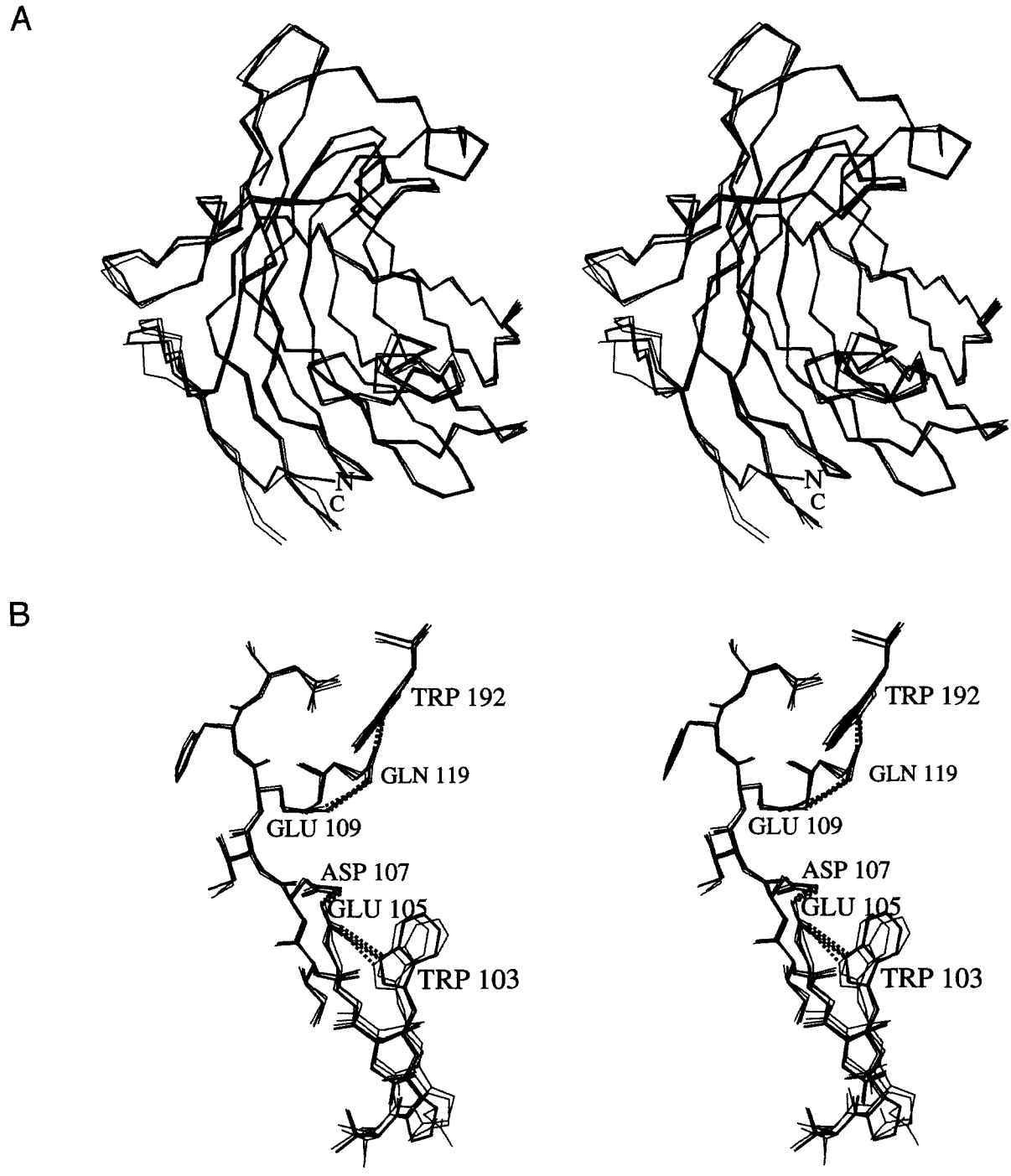

Fig. 3. Least squares superposition of the α-carbon positions of BGLL with BGLM (both molecules of the asymmetric unit) and H(A16-M). (a) α-Carbon trace of the polypeptide chains, (b) close-up into the active site showing the catalytic-site residues and the conserved hydrogen bonding (dashed lines) between them. BGLL is drawn in black, the other models in grey.

3.2. Comparison with known 1,3-1,4-β-glucanases

The sequence identity is 75.5% between BGLL and BGLM and 77.1% between BGLL and H(A16-M). Superpositions of the C_α positions of BGLL with those of BGLM result in rms differences of 1.11 Å and 1.15 Å for the two molecules in the asymmetric unit of BGLM, respectively. The superposition with H(A16-M), if all C_α positions are considered, gives an rms deviation of 0.68 Å. If only the *B. macerans* derived residues from 17 to 214 of H(A16-M) are used in the fit it reduces to 0.41 Å. In Fig. 3a the H(A16-M) model and the two models of BGLM are superimposed onto BGLL. As the structure of H(A16-M) was refined at 1.16 Å resolution it is more suitable for comparison with the 1.8 Å structure of BGLL than the BGLM models. In spite of the fact that 13 of the 16 N-terminal amino acids are identical in BGLL and H(A16-M), structural differences are largest in this region. This is explained by the packing of the molecules: in H(A16-M) amino acids Gln-1, Tyr-13 and Asn-14 are involved in crystal contacts, whereas in

Table 2
Intermolecular hydrogen bonds

Amino acid	Atom	Amino acid	Atom	Translation			Distance (Å)
Met-41	N	Glu-191	$O_{\varepsilon 2}$	0	−1	0	2.8
Ser-25	O	Arg-210	$N_{\eta 1}$	−1	0	0	2.8
Asn-26	O	Arg-210	$N_{\eta 2}$	−1	0	0	3.0
Glu-46	$O_{\varepsilon 2}$	Lys-58	N_{ζ}	1	0	0	2.8
Ser-53	O_{γ}	Asn-154	$N_{\delta 2}$	−1	0	−1	3.0
Tyr-56	O_{η}	Gln-171	N	−1	0	−1	3.0
Asn-57	N	Asn-154	$O_{\delta 1}$	−1	0	−1	2.7

BGLL none of the first 16 residues is involved in crystal packing. Therefore the structural differences in the first loop at the N-terminus are caused by packing in the H(A16-M) crystals and not by differences in the sequences. BGLL seems to be nearly unaffected by crystal contacts in general, as the mean distance of those C_{α} atoms involved in contacts from their counterparts in H(A16-M) is close to the mean distance for all C_{α}-atoms (0.42 Å compared to 0.41 Å).

3.3. Active site

Amino acid differences between the substrate binding channels of BGLL and H(A16-M) are restricted to Met-29 (H(A16-M), Val) and Asp-99 (H(A16-M), His). This is not expected to cause a drastic change in the carbohydrate binding behavior. The His-99 imidazole forms a hydrogen bond to a glucose hydroxyl group according to the crystal structure of H(A16-M) in a covalent complex with a cellobioside inhibitor [20]. Asp-99 in BGLL may assume this function, possibly with reversed polarity. The catalytic-site geometry is nearly identical in all models obtained so far (Fig. 3b). Therefore, the stereochemistry of the reaction as determined for BGLL [15] may be expected to be the same in both BGLM and H(A16-M).

We summarize that our studies of the structures of bacterial 1,3-1,4-β-glucanases show the tertiary structures of these enzymes to be remarkably similar. Not only the wildtype and hybrid enzymes BGLL, BGLM and H(A16-M), but also variants with new chain termini caused by a circular permutation of their genes [9], fold to this compact and stable protein domain in spite of different crystallization conditions and packing with little conservation of crystal contacts in a number of different space groups, P1, $P2_1$, C2 and $P2_12_12_1$. This observation leads to the conclusions that the β-glucanase jellyroll fold is of pronounced rigidity and that all homologous bacterial 1,3-1,4-β-glucanases share the same enzyme mechanism.

Acknowledgements: X-ray diffraction data used in this study were collected in the laboratory of Prof. W. Saenger (Freie Universität Berlin, Germany). Supported by the Deutsche Forschungsgemeinschaft (He 1318/9) and the Fonds der Chemischen Industrie.

References

[1] Murphy, N., McConnell, D.J. and Cantwell, B.A. (1984) Nucleic Acids Res. 12, 5355–5367.
[2] Hofemeister, J., Kurtz, A., Borriss, R. and Knowles, J. (1986) Gene 49, 177–187.
[3] Borriss, R., Büttner, K. and Mäntsälä, P. (1990) Mol. Gen. Genet. 222, 278–283.
[4] Bueno, A., Vazquez de Aldana, C.R., Correa, J., Villa, T.G. and del Rey, F. (1990) J. Bacteriol. 172, 2160–2167.
[5] Gosalbes, M.J., Perez-Gonzalez, A., Gonzalez, R. and Navarro, A. (1991) J. Bacteriol. 173, 7705–7710.
[6] Lloberas, J., Perez-Pons, J.A. and Querol, E. (1991) Eur. J. Biochem. 197, 337–343.
[7] Louw, M.E., Reid, S.J. and Watson, T.G. (1993) Appl. Microbiol. Biotechnol. 38, 507–513.
[8] Fincher, G.B., Lock, P.A., Morgan, M.M., Lingelbach, K., Wettenhall, R.E.H., Mercer, J.F.B., Brandt, A. and Thomson, K.K. (1986) Proc. Natl. Acad. Sci. USA 83, 2081–2085.
[9] Hahn, M., Piotukh, K., Borriss, R. and Heinemann, U. (1994) Proc. Natl. Acad. Sci. USA 91, 10417–10421.
[10] Hahn, M., Olsen, O., Politz, O., Borriss, R. and Heinemann, U. (1995) J. Biol. Chem. 270, 3081–3088.
[11] Hahn, M., Keitel, T. and Heinemann, U. (1995) Eur. J. Biochem., in press.
[12] Planas, A., Juncosa, M., Lloberas, J. and Querol, E. (1992) FEBS Lett. 308, 141–145.
[13] Juncosa, M., Pons, J., Dot, T., Querol, E. and Planas, A. (1994) J. Biol. Chem. 269, 14530–14535.
[14] McCarter, J.D. and Withers, S.G. (1994) Current Opinion Struct. Biol. 4, 885–892.
[15] Malet, C., Jimenez-Barbero, J., Bernabé, M., Brosa, C. and Planas, A. (1993) Biochem. J. 296, 753–758.
[16] Messerschmidt, A. and Pflugrath, J.W. (1987) J. Appl. Cryst. 20, 306–315.
[17] Leslie, A.G.W. (1992) Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 26.
[18] Sheldrick, G.M. (1976) SHELX76. Program for Crystal Structure Determination. University of Cambridge, England.
[19] Brünger, A.T. (1992) X-PLOR Manual, Version 3.1, Yale University Press, New Haven.
[20] Keitel, T., Simon, O., Borriss, R. and Heinemann, U. (1993) Proc. Natl. Acad. Sci. USA 90, 5287–5291.
[21] Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M. (1991) Acta Crystallogr. A47, 110–119.
[22] Tronrud, D.E., Ten Eyck, L.F. and Matthews, B.W. (1987) Acta Crystallogr. A43, 489–500.
[23] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–291.
[24] Brünger, A.T. (1992) Nature 355, 472–474.
[25] Hutchinson, E.G. and Thornton, J.M. (1990) Proteins: Struct. Funct. Genet. 8, 203–212.
[26] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.
[27] Pons, J., Planas, A. and Querol, E. (1995) Protein Engineering, in press.
[28] Evans, S.V. (1993) J. Mol. Graphics 11, 134–138.
[29] Kraulis, J.P. (1991) J. Appl. Crystallogr. 24, 946–950.