Biochemistry

 Volume 39, Number 17 May 2, 2000

Articles

Comparative Structural Analysis and Substrate Specificity Engineering of the Hyperthermostable β-Glucosidase CelB from *Pyrococcus furiosus*†

Thijs Kaper,*,‡ Joyce H. G. Lebbink,‡ Jeroen Pouwels,‡ Jürgen Kopp,§ Georg E. Schulz,§ John van der Oost,‡ and Willem M. de Vos‡

Laboratory of Microbiology, Department of Biomolecular Sciences, Wageningen University, Hesselink van Suchtelenweg 4, NL-6703 CT Wageningen, The Netherlands, and Institut für Organische Chemie und Biochemie, Albert-Ludwigs-Universität, Albertstrasse 21, D-79104 Freiburg im Breisgau, Germany

Received October 22, 1999; Revised Manuscript Received January 28, 2000

ABSTRACT: The substrate specificity of the β-glucosidase (CelB) from the hyperthermophilic archaeon *Pyrococcus furiosus*, a family 1 glycosyl hydrolase, has been studied at a molecular level. Following crystallization and X-ray diffraction of this enzyme, a 3.3 Å resolution structural model has been obtained by molecular replacement. CelB shows a homo-tetramer configuration, with subunits having a typical $(\beta\alpha)_8$-barrel fold. Its active site has been compared to the one of the previously determined 6-phospho-β-glycosidase (LacG) from the mesophilic bacterium *Lactococcus lactis*. The overall design of the substrate binding pocket is very well conserved, with the exception of three residues that have been identified as a phosphate binding site in LacG. To verify the structural model and alter its substrate specificity, these three residues have been introduced at the corresponding positions in CelB (E417S, M424K, F426Y) in different combinations: single, double, and triple mutants. Characterization of the purified mutant CelB enzyme revealed that F426Y resulted in an increased affinity for galactosides, whereas M424K gave rise to a shifted pH optimum (from 5.0 to 6.0). Analysis of E417S revealed a 5-fold and a 3-fold increase of the efficiency of hydrolyzing *o*-nitrophenol-β-D-galactopyranoside-6-phosphate, in the single and triple mutants, respectively. In contrast, their activity on nonphosphorylated sugars was largely reduced (30−300-fold). The residue at position E417 in CelB seems to be the determining factor for the difference in substrate specificity between the two types of family 1 glycosidases.

Glycosyl hydrolases are widely spread among living organisms. Members of this enzyme superfamily are involved in a broad range of physiological processes, such as food supply and storage, cell wall synthesis and degradation, defense systems, and signaling events (*1*). A large number of glycosyl hydrolases have been isolated and characterized, and the corresponding amino acid sequences have been determined. At the moment, comparison of these sequences has led to the identification of 77 families of glycosyl hydrolases (*2*). Family 1 consists of β-glycosidases, 6-phospho-β-glycosidases, lactase−phlorizin hydrolases, and myrosinases originating from all three domains of life: Bacteria, Archaea, and Eukarya. These enzymes hydrolyze their substrate with retention of configuration at the anomeric

† This work has been partly supported by Contracts FAIR CT96-1048 and BIO 4-CT96-0488 of the European Union.

* To whom correspondence should be addressed. Tel: 31-317-483110, fax: 31-317-483829, e-mail: thijs.kaper@algemeen.micr.wau.nl.

‡ Wageningen University.

§ Albert-Ludwigs-Universität.

10.1021/bi992463r CCC: $19.00
Published on Web 04/07/2000

carbon atom, which occurs via a double displacement mechanism, in which a covalent intermediate is believed to be formed (*3*−*5*). Two glutamate residues serve as a general acid/base and a nucleophile in this reaction. Although family 1 glycohydrolases in general display a broad substrate specificity, the glycosidases have been optimized for the hydrolysis of nonphosphorylated glycosides and have hardly any activity on phosphorylated sugars, while the opposite is true for the family 1 6-P-glycosidases.

In the last years, the 3D structures have been solved of various family 1 glycosyl hydrolases from Eukarya (*6*, *7*), Bacteria (*8*−*10*), and Archaea (*11*, *12*). The enzymes display a common $(\beta\alpha)_8$-barrel motif, and, being members of the glycosyl hydrolase 4/7-superfamily, all have the two catalytic glutamates located at the C-terminal end of β-strands 4 and 7 (*13*). The involvement of the two glutamates in catalysis has been well established by mutagenesis and inhibitor studies (*14*−*18*). Recently, the 3D structure of an inactive mutant of the *Lactococcus lactis* 6-P-galactosidase LacG has been elucidated with a galactose-6-P bound in the active site (*10*).

Previously, a β-glucosidase has been purified from the hyperthermophilic archaeon *Pyrococcus furiosus* (*19*). The enzyme was active as a homo-tetramer with 58 kDa subunits and displayed a high thermostability, with a half-life of thermal inactivation of 85 h at 100 °C. The corresponding *celB* gene has been cloned and functionally expressed in *E. coli*, and its sequence characterization identified the *P. furiosus* β-glucosidase as a family 1 glycosyl hydrolase, in which Glu372 was identified as the nucleophilic catalytic residue by site-directed mutagenesis (*16*). A translational fusion of the *celB* gene to the T7 promoter of the pET9d vector, in which an alanine residue was inserted after the N-terminal methionine, allowed for stable overproduction in *E. coli* of the *P. furiosus* β-glucosidase (*20*). The activity, stability, and kinetic parameters of the resulting enzyme, referred to here as CelB, were comparable to those of the β-glucosidase isolated from *P. furiosus* (*21*). CelB serves as a model system to study the substrate specificity of glycosyl hydrolase family 1, since it has been well characterized, and can be produced in *E. coli* in wild-type and mutant forms and displays extreme stability, broad specificity, and high catalytic activity (*19*, *20*, *22*−*24*). In addition, CelB has developed into a model enzyme that may have wide applications, such as disaccharide hydrolysis and synthesis of glycoconjugates and oligosaccharide production (*25*−*27*).

Here, we report the investigation and mutagenesis of the active site of CelB. Using X-ray diffraction data, a 3D model of the *P. furiosus* β-glucosidase was built by the molecular replacement method, in which the structures of *L. lactis* 6-P-β-galactosidase LacG (*8*) and *Sulfolobus solfataricus* β-glycosidase LacS (*11*) served as search models. Based on the structural model, we designed three substitutions at three positions to adapt the enzyme to the hydrolysis of phosphorylated galactosides. CelB mutants have been purified, characterized, and compared to the wild-type enzyme. The results show that a single substitution is mainly responsible for a significant increase of the catalytic activity on phosphorylated galactose.

MATERIALS AND METHODS

Strains, Plasmids, and Chemicals. *Escherichia coli* BL21-(DE3) (*hsdS gal* λ*c*I*ts* 857 *ind*1 *S*am7 *nin*5 *lac*UV5-T7 gene 1) or *E. coli* MC1061was used for heterologous expression; *E. coli* TG1 (*supE hsd*Δ5 *thi* Δ(*lac-proAB*) F′[*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]) was used in the construction of derivatives of pET9d (T7 promoter, kan^R) (Novagen) following established procedures (*28*). Plasmid pLUW511 is a pET9d derivative carrying the *P. furiosus celB* gene (GenBank accession no. AF013169), derived from pLUW510 (*16*), translationally fused to the bacteriophage T7 ϕ10 gene start codon (*20*).

All chemicals used were of analytical grade. *O*-Nitrophenyl-β-D-glucopyranoside (oNp-glu),[1] *o*-nitrophenyl-β-D-galactopyranoside (oNp-gal), and *o*-nitrophenyl-β-D-galactopyranoside-6-phosphate (oNp-gal-6-P) were purchased from Sigma (Zwijndrecht, The Netherlands). *Pfu* DNA polymerase was obtained from Stratagene.

Construction of Mutants. Mutations were introduced in *celB* using *Pfu* polymerase in the PCR-based overlap extension method (*29*). For each mutation, a sense/antisense primer pair was designed. BG266 (5′-AACAGACAACTAC**TCC**TG-3′)/BG267 (5′-GGGCCCA**GGA**GTAGTTGTCTGTT-3′) introduced mutation E417S (introduced mutations in boldface, introduced restriction sites underlined, and restriction enzymes in parentheses). BG268 (5′-CAAGGGTTCAGGA**AA**AGAT**AC**GGATTGGTTT-3′)/BG269 (5′-AAACCAATCC**GT**ATCT**TT**CCTGAACCCTTG-3′) introduced M424K/F426Y. BG333 (5′-GGGTTCAGGAA**AA**AGATT**C**GGATTGGTT-3′)/BG334 (5′-AACCAATCC**G**AATCT**TT**CCTGAACCC-3′) (*Xmn*I) introduced M424K. F426Y was introduced by BG335 (5′-GGTTCAGGATGAG**GT**A**C**GGATTGGTT-3′)/BG336 (5′-AACCAATCC**GT**A**C**CTCATCCTGAACC-3′) (*Rsa*I). In each case the sense primer BG238 (5′-GCGCGCC*ATG*GCAAAGTTCCCAAAAAACTTCATGTTTG-3′, *celB* start codon in italics) (*Nco*I) and antisense primer BG239 (5′-CGCGCGGATCC*CTA*CTTTCTTGTAACAAATTTGAGG-3′, *celB* stop codon in italics) (*Bam*HI) were used as flanking primers for the amplification of the *celB* gene. The resulting PCR products were digested with *Nco*I and *Bam*HI and ligated in *Nco*I/*Bam*HI-digested pET9d. Plasmid pLUW511 was used as a template in PCR reactions for the construction of pLUW513 (E417S), 515 (M424K), and 522 (F426Y). Plasmid pLUW513 was used as a template in the construction of pLUW519 (E417S/M424K/F426Y). Plasmid pLUW518 (M424K/F426Y) was constructed by replacing a pLUW519 *Pst*I−*Afl*II fragment with a pLUW511 *Pst*I−*Afl*II fragment. Mutations were verified by DNA sequence analysis using the Thermosequenase cycle sequencing kit with infrared dye labeled primers. Reactions were analyzed on a Licor 4000L automated sequencer (Licor, Lincoln, NE) (data not shown).

Enzyme Production and Purification. CelB was routinely produced in *E. coli* BL21(DE3) harboring pET9d derivatives and purified from heat-incubated lysates following anion exchange and gel filtration chromatography (*20*). This resulted in pure CelB, as was confirmed by SDS−PAGE

[1] Abbreviations: oNp-glu, *o*-nitrophenol-β-1,4-D-glucopyranoside; oNp-gal, *o*-nitrophenol-β-1,4-D-galactopyranoside; oNp-gal-6-P, *o*-nitrophenol-β-1,4-D-galactopyranoside-6-phosphate; pNp-gluc, *p*-nitrophenol-β-1,4-D-glucuronate.

electrophoresis, and was stored at 4 °C with 0.02% sodium azide. Protein concentrations were determined spectrophotometrically at 280 nm with a molar extinction coefficient ϵ_{280} of 128 280 M^{-1} cm^{-1}, calculated according to Gill and Von Hippel (*30*).

Enzyme Preparation, Crystallization, and Data Collection. *P. furiosus* β-glucosidase was purified from *E. coli* MC1061 cells harboring pLUW510 as essentially described previously except for the heat incubation of 15 min at 60 °C (*16*). Initial crystallization conditions were determined after screening according to Jancarik and Kim (*31*). After fine-tuning, the enzyme was found to crystallize optimally by vapor diffusion using the hanging-drop method at 20 °C. Drops of 10 μL contained 10 mg/mL protein, 8% PEG-400, 150 mM $CaCl_2$, and 10 mM Tris, pH 8.0, and the reservoir contained 20% PEG-400. Within a few days, crystals grew to an average size of 600 × 600 × 600 μm^3. The crystals diffracted merely to medium resolution. A native data set to 3.3 Å could be collected at RT with a multiwire area detector (X1000, Nicolet/Siemens) on a rotating anode with exposure times up to 240 s per frame. The space group was I_132 with a unit cell of $a = b = c = 219.3$ Å. The structures of 6-phospho-β-galactosidase LacG of *L. lactis* (PDB entry 1PBG) (*8*) and β-glycosidase LacS of *S. solfataricus* (PDB entry 1GOW) (*11*) were used as search models for molecular replacement and subsequent construction of the structural model of *P. furiosus* β-glucosidase. Solutions for *L. lactis* LacG and *S. solfataricus* LacS found with the programs AMORE (*32*) and AUTOAMORE (*32*), respectively, were consistent with each other. A subunit of LacS was used for an initial rigid body refinement with X-plor (*33*). After introduction of all CelB residues, the structure was refined with X-plor (*33*) using bulk solvent correction. The *R*-factor dropped to 32.5% (R_{free} = 39.0%), and the common *B*-factor was 31 $Å^2$. All residues were in the energetically most favorable or allowed regions of the Ramachandran plot of backbone dihedral angles.

Kinetic Analyses. *(A) pH Optima.* The pH optima of wild-type CelB and CelB mutants were determined at 90 °C in combined citrate/phosphate buffer (80 mM/80 mM) in a pH range of 3.5−8.0 (set at 25 °C). As substrates, oNp-glu, oNp-gal, and oNp-gal-6-P were used. In the case of oNp-gal-6P, reaction vials (1.5 mL) containing 49 μL of 40 mM oNp-gal-6-P in buffer were preincubated in a water bath at 90 °C for 2 min. The reaction was started by the addition of 1 μL of enzyme solution to the vials. After 5 min, the reaction was stopped by the addition of 100 μL of ice-cold 0.5 M Na_2CO_3, which terminated the reaction and augmented the molar extinction coefficient of the released nitrophenol; 125 μL samples were transferred to a 96-well microtiter plate, and the extinction of the reaction mixtures was determined at 405 nm in a microtiter plate reader (SLT Lab Instruments 340 ATTC). Maximum extinction was set at 100%. When oNp-glu and oNp-gal were used as substrates, reaction vials containing 195 μL of 10 mM substrate in buffer were treated as for oNp-gal-6-P; only the reactions were initiated by the addition of 5 μL of enzyme and stopped with 0.4 mL of carbonate, and 150 μL samples were transferred to a microtiter plate for analysis.

(B) K_m and V_{max} Determinations. The K_m and V_{max} of wild-type CelB and CelB mutants were measured for the hydrolysis of the three substrates at 90 °C in 100 mM citrate/phosphate buffer. We followed the release of nitrophenol at 405 nm (pH 5.0) or 470 nm (pH 6.0) in quartz cuvettes using a spectrophotometer equipped with a temperature controller (Hitachi, San Jose, CA). For each substrate, the activities were measured at differerent concentrations (oNp-glu: 0.05−50 mM; oNp-gal: 0.10−50 mM; oNp-gal-6-P: 0.4−24 mM). The reaction volume was 1.00 mL for oNp-glu and oNp-gal and 0.50 mL for oNp-gal-6-P. In each case, an enzyme volume of 5.0 μL was added to start the reaction. Molar extinction coefficients of *o*-nitrophenol at 90 °C in assay mixtures were the following: $\epsilon_{405;pH5.0}$, 0.6243 cm^{-1} mM^{-1}; $\epsilon_{470;pH6.0}$, 0.6339 cm^{-1} mM^{-1}; $\epsilon_{405,pH4.5}$, 0.5614 cm^{-1} mM^{-1}; $\epsilon_{405,pH4.0}$, 0.5398 cm^{-1} mM^{-1}. K_m and V_{max} values were calculated by fitting the activities according to Michaelis−Menten kinetics using the nonlinear regression program TableCurve (Jandel Scientific, AISN Software).

RESULTS

CelB Structure. The *P. furiosus* β-glucosidase CelB is the most thermostable and thermoactive representative of family 1 glycosyl hydrolases and shows a broad substrate specificity. To characterize its structure−function relation, the *P. furiosus* β-glucosidase was purified from an overproducing *E. coli* strain to homogeneity, crystallized by the hanging-drop method, and subjected to X-ray diffraction. A dataset was collected at 3 Å resolution, which allowed for the analysis of its tertiary structure and active site. The tetrameric structure of the enzyme was established in the molecular replacement procedure (Figure 1). The search models in this procedure were the 3D structures of the *L. lactis* 6-phospho-β-galactosidase LacG and the *S. solfataricus* β-glycosidase LacS, which share 16% and 53% amino acid identity with the pyrococcal enzyme, respectively. Attempts were made to obtain more regular crystals from CelB variants with altered surface residues, but crystals that diffracted at a higher resolution were not obtained (data not shown). Therefore, the initial structural model was used in the analysis of the CelB tertiary structure, the active site, and for the design of specific substitutions.

The CelB homo-tetramer has point symmetry *222*, with one subunit in the crystallographically asymmetric unit. The enzyme consists of two dimers, of which the two barrels are placed next to each other. In the final tetramer, two dimers appear to be placed on top of each other with the barrel openings facing each other (Figure 1). The overall size of the tetramer is 110 Å × 90 Å × 55 Å, and its appearance is that of a slightly twisted square. The axes of the $(\beta\alpha)_8$-barrels are almost parallel to one of the tetramer axes (Figure 1). Each subunit of CelB consists of a single domain of 472 amino acids, with 18 helices and 16 β-strands. The center of the monomer is formed by an $(\beta\alpha)_8$-barrel, with the catalytic residues Glu207 and Glu372 at the C-terminal end of the 4th and 7th β-strand of the barrel, respectively.

Architecture of the CelB Active Site. The positions of the active site residues in LacG and CelB are very well conserved, as illustrated by a superposition (Figure 2). Asn17, Arg77, His150, Asn206, Tyr307, and Trp410 are conserved residues and likely to be involved in substrate binding (*10*). Glu207 and Glu372 in CelB are the equivalents of the catalytic glutamate residues in LacG. The importance of

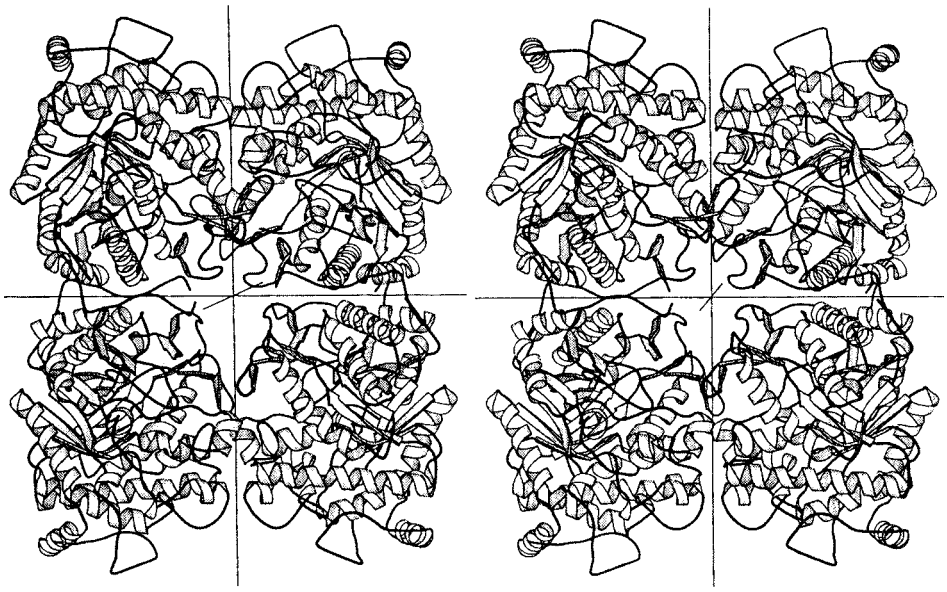

FIGURE 1: Ribbon model of *P. furiosus* β-glucosidase, viewed along one of the 2-fold axes of the tetramer.

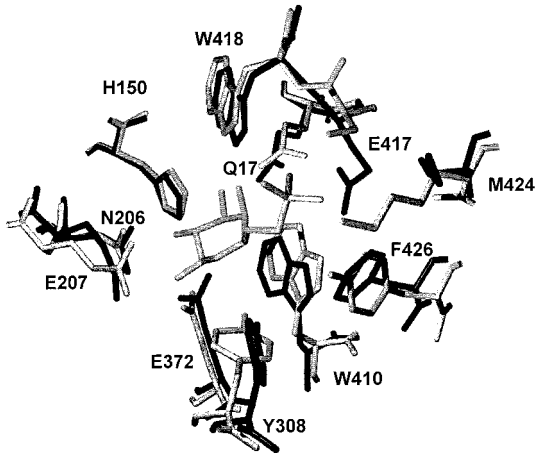

FIGURE 2: Superposition of *P. furiosus* CelB (black) with *L. lactis* LacG (gray) that has galactose-6-phosphate bound in the active site. The residues that are involved in the glucose or galactose binding in both enzymes have been represented. The residues have been indicated according to CelB numbering. The calculated RMS deviation of equivalent atoms of these residues is about 1 Å, which corresponds to the expected accuracy of the CelB model. The general superposition was based on the backbone atoms of both proteins and was performed using the Swiss PDB Viewer (*42*). The figure was generated using the program POV-Ray (Hallam Oaks Pty. Ltd.).

```
Pfu_CelB   NAYELP--MIITENGMADAADR----------YRPHYLVSHLKA : 393
Sso_LacS   R-YHLY--MYVTENGIADDADY----------QRPYYLVSHVYQ : 408

Lla_LacG   D-YPNYKKIYITENGLGYKDEFVDN--TVYDDGRIDYVKQHLEV : 405
Sau_LacG   D-YPNYHKIYITENGLGYKDEFIESEKTVHDDARIDYVRQHLNV : 407
Eco_BglB   R-YQ--KPLFIVENGLGAKDS-VEADGSIQDDYRIAYLNDHLVQ : 391
                    #

Pfu_CelB   VYNAMKEGADVRGYLHWSLTDNYEWAQG-FRNRFGLVYVDFETK : 436
Sso_LacS   VHRAINSGADVRGYLHWSLADNYEWASG-FSNRFGLLKVDYNTK : 451

Lla_LacG   LSDAIADGANVKGYFIWSLMDVFSWSNG-YEKRYGLFYVDFDTQ : 448
Sau_LacG   IADAIIDGANVKGYFIWSLMDVFSWSNG-YEKRYGLFYVDFETQ : 450
Eco_BglB   VNEAIADGVDIMGYTSWGPIDLVSASHSQMSKRYGFIYVDRDDN : 435
                           *         * *

Pfu_CelB   K-----RYLRPSALVFREIATQKEIPEELAHLADLKFVT--RK. : 472
Sso_LacS   R-----LYWRPSALVYREIATNGAITDEIEHLNSVPPVKPLRH. : 489

Lla_LacG   E-----RYPKKSAHWYKKLAETQVIE.                  : 468
Sau_LacG   E-----RYPKKSAYWYKELAETKEIK.                  : 470
Eco_BglB   GEGSLTRTRKKSFGWYAEVIKTRGLSLKKITIKAP.         : 470
```

FIGURE 3: Alignment of partial amino acid sequences. *Pfu* CelB, *P. furiosus* β-glucosidase CelB (Swiss Prot Q51763); *Sso* LacS, *S. solfataricus* β-glycosidase LacS (P22498); *Lla* LacG, *L. lactis* 6-phospho-β-galactosidase LacG (P11546); *Sau* LacG, *Staphylococcus aureus* 6-phospho-β-galactosidase LacG (P11175); *Eco* BglB, *Escherichia coli* 6-phospho-β-glucosidase BglB (P11988). Conserved residues in all proteins have been shaded light gray; residues that interact with the phosphate group of galactose-6-phosphate in *L. lactis* LacG have been indicated (*) and shaded black in the 6-phospho-β-glycosidases and dark gray at the corresponding positions in the β-glycosidases. The conserved catalytic nucleophile in all enzymes has been indicated (#).

Glu372 in catalysis at 100 °C has been shown by site-directed mutagenesis (*16*). The average distance between the oxygen atoms of these glutamate carboxylic groups is 4.3 Å ($\pm$ 1Å) in CelB, which is in the range of the general observed distance in retaining glycosyl hydrolases (~5 Å) (*34*). The active site is located in the center of a subunit and can be reached from the outside by a 20 Å long channel, which is located between the fifth and sixth $\beta\alpha$-motives and is opened by a 50° bending of the helix of the fifth $\beta\alpha$-unit.

Design and Production of Wild-Type and Mutant P. furiosus CelB. In 6-P-β-galactosidase LacG of *Lactococcus lactis*, Ser428, Lys435, and Tyr437 form a binding site for the phosphate group of galactose-6-phosphate (*10*). These three residues are highly conserved among family 1 6-P-β-galactosidases (Figure 3). In the *P. furiosus* CelB, the corresponding residues at these positions are Glu417, Met424, and Phe426, residues which, in turn, are highly conserved among family 1 β-glycosidases that hydrolyze nonphosphorylated substrates. To study their contribution to substrate specificity, the residues that create the phosphate binding site in LacG were introduced in CelB by a PCR-based method. This resulted in three CelB single mutants (E417S, M424K, and F426Y), one double mutant (M424K/F426Y), and one triple mutant (E417S/M424K/F426Y). The wild-type and mutated *celB* genes were overexpressed in *E. coli*, all resulting in significant overproduction of active enzyme, which amounted up to more than 20% of the total cell protein. Subsequently, the mutant CelB enzymes were purified to apparent homogeneity as judged by SDS−PAGE

Table 1: pH Optima of Wild-Type CelB and CelB Mutants at 90 °C in 0.1 M Sodium Citrate/NaP_i

	oNp-glu	oNp-gal	oNp-gal-6-P
CelB wt	5.0	5.0	4.0
CelB E417S	4.5	4.5	5.0
CelB M424K	6.0	5.75	4.5
CelB F426Y	5.0	5.0	4.0
CelB double mutant	6.0	6.25	4.5
CelB triple mutant	5.0	5.0	6.0

and UV spectroscopy (not shown) and used for further characterization and comparison with the wild-type CelB.

pH Optima of Mutant and Wild-Type CelB. Wild-type CelB hydrolyzes the chromogenic substrate pNp-β-glu optimally at pH 5.0 (*19*). The glycosidic bond is cleaved by a common double displacement mechanism (*24*). The pH optimum for this mechanism has been proposed to be determined by the pK_a values of the two catalytic residues (*35*). Since the substitutions E417S and M424K introduce charge differences in the active site, these might influence the pK_a values of the catalytic glutamate residues and, as such, change the pH optimum for hydrolysis. Therefore, the pH optima for the hydrolysis of *o*-nitrophenol-β-D-glucopyranoside (oNp-glu), *o*-nitrophenol-β-D-galactopyranoside (oNp-gal), and *o*-nitrophenol-β-D-galactopyranoside-6-phosphate (oNp-gal-6-P) were determined for different mutants (Table 1). The stability of the mutants at the different pH values was found to be equal to that of wild-type CelB, with the exception of CelB M424K and CelB M424K/F426Y, which were more sensitive to inactivation below pH 4.5 (not shown).

The pH optima of wild-type and mutant CelB enzymes for the hydrolysis of oNp-glu and oNp-gal were essentially the same (Table 1). However, large differences were observed in the shape of the pH curves. Three groups could be distinguished: (i) wild-type CelB, M424K, and F426Y that showed an optimum activity at a broad pH range for oNp-glu and oNp-gal, but a narrow range for oNp-gal-6-P with a lower optimal pH than with oNp-glu and oNp-gal as substrates (Figure 4A,B); (ii) E417S and the triple mutant that showed a pH dependency of activity for the three substrates that was opposite of that of wild-type CelB (Figure 4C,D); (iii) the double mutant M424K/F426Y hydrolyzed both oNp-glu and oNp-gal as well as oNp-gal-6-P optimally at a relatively wide pH range, the optimal pH for hydrolysis with oNp-gal-6-P being lower than with oNp-glu and oNp-gal as substrates (Figure 4C,D). The mutation F426Y did not have any effect on the pH optimum of hydrolysis, compared to the wild type. However, M424K hydrolyzed oNp-glu and oNp-gal optimally at pH 6.0, 1 unit higher than the wild-type. This mutant also showed the fastest hydrolysis of oNp-gal-6-P at a higher pH than the wild-type.

Kinetic Parameters of Mutant and Wild-Type CelB. The kinetic constants at optimal pH for the hydrolysis of oNp-glu and oNp-gal were determined (Table 2), as well as those for the hydrolysis of oNp-gal-6-P (Table 3). The wild-type CelB shows high affinity for oNp-glu (K_M value of 0.28 mM) and a lower affinity for oNp-gal (K_M value of 2.3 mM). All introduced mutations decreased the affinity for oNp-glu compared to the wild type. Moreover, the turnover number decreased for most mutants, resulting in lower efficiencies

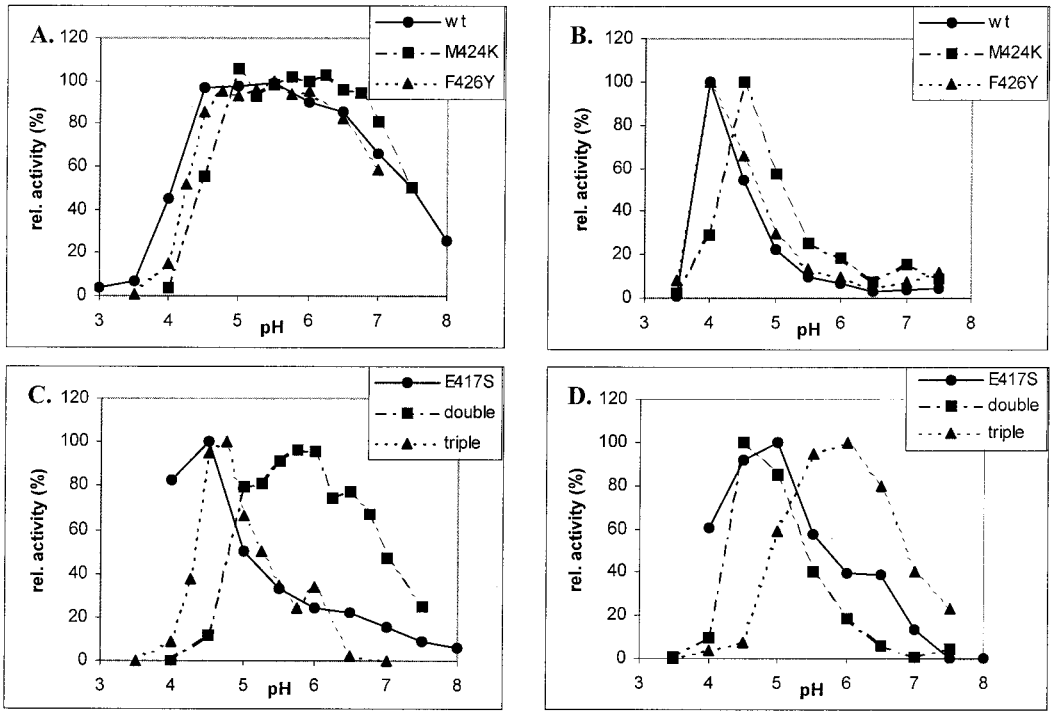

FIGURE 4: Typical effects of pH on enzyme activity at 90 °C in combined citrate−phosphate buffer. Assays were performed with 8.3 ng to 3.3 μg of enzyme. Maximum activity in each case was set at 100%. (A) Activity on oNp-glu of wild-type CelB (max. 920 units/mg), CelB M424K (max. 805 units/mg), and CelB F426Y (max. 910 units/mg); (B) activity on oNp-gal-6-P of wild-type CelB (max. 31 units/mg), CelB M424K (max. 5 units/mg), and CelB F426Y (max. 14 units/mg); (C) activity on oNp-glu of CelB E417S (max. 29 units/mg), CelB double mutant (max. 1240 units/mg), and CelB triple mutant (max. 124 units/mg); (D) activity on oNp-gal-6-P of CelB E417S (max. 89 units/mg), CelB double mutant (max. 27 units/mg), and CelB triple mutant (max. 54 units/mg). 1 Unit corresponds to the release of 1 μmol oNp per minute.

Table 2: Kinetic Constants of Wild-Type and Mutant CelB with oNp-glu and oNp-gal as Substrates at 90 °C in 0.1 M Sodium Citrate/NaP_i at Determined pH Optimum[a]

	oNp-glu			oNp-gal		
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
CelB wt	0.28 ± 0.03	860 ± 25	4048	2.3 ± 0.4	2050 ± 86	899
CelB E417S	29 ± 10	3300 ± 630	113	42 ± 15	1060 ± 250	25.3
CelB M424K	0.51 ± 0.12	772 ± 47	1510	4.8 ± 0.3	1880 ± 50	396
CelB F426Y	0.57 ± 0.09	840 ± 37	1484	0.89 ± 0.05	892 ± 13	1000
CelB double	3.0 ± 0.5	1400 ± 89	422	3.0 ± 0.6	969 ± 65	325
CelB triple	59 ± 7	777 ± 65	13	40 ± 8	175 ± 21	4.4

[a] Amounts of 83 ng of CelB wt, 0.11 μg of CelB E417S, 0.13 μg of CelB M424K, 0.23 μg of CelB F426Y, 0.14 μg of CelB double mutant, and 1.3 μg of CelB triple mutant were used in the assays.

Table 3: Kinetic Constants of Wild-Type and Mutant CelB at 90 °C with oNp-gal-6-P as Substrate in 0.1 M Sodium Citrate/NaP_i at pH Optimum[a]

	oNp-gal-6-P		
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
CelB wt	31 ± 6	49 ± 6	1.6
CelB E417S	22 ± 2	171 ± 11	7.7
CelB M424K	−	(11)[b]	−
CelB F426Y	38 ± 10	26 ± 5	0.7
CelB double mutant	−	(58)[b]	−
CelB triple mutant	16 ± 2	69 ± 5	4.2
LacG[c]	0.4	83	208

[a] Amounts of 3.3 μg of CelB wt, 2.1 μg of CelB E417S, 26 μg of CelB M424K, 47 μg of CelB F426Y, 2.9 μg of CelB double mutant, and 3.8 μg of CelB triple mutant were used in the assays. [b] Data could not be fitted according to Michaelis−Menten kinetics; highest measured k_{cat} shown. [c] LacG activity measured at 25 °C; data obtained from (*41*).

for the hydrolysis of oNp-glu. The affinity for oNp-gal decreased in E417S, M424K, and the triple mutant, was unaffected in the double mutant, but was significantly increased in F426Y. This made F426Y slightly more efficient on oNp-gal than the wild type. The K_M values for oNp-glu and oNp-gal in the mutants that contain the mutation E417S (E417S and the triple mutant) increased 20-fold to 200-fold, but due to the limited solubility of oNp-glu and oNp-gal, it was not possible to perform assays at saturating conditions for these mutants.

Although wild-type CelB is able to hydrolyze oNp-gal-6P, its efficiency is low with a K_M of 31 mM and a k_{cat} of 49 s^{-1} under optimal conditions (Table 3). Remarkably, mutant F426Y has a slightly higher K_M value and a reduced k_{cat} value with respect to wild-type CelB, making it even less efficient in hydrolysis of the phosphorylated substrate. The CelB mutants that contain the substitution E417S had lost their affinity for nonphosphorylated glycosides, but these variants (mutant E417S and the triple mutant) have increased affinities and turnover numbers for the hydrolysis of oNp-gal-6-P. As a consequence, E417S and the triple mutant are 4.9-fold and 2.7-fold more efficient, respectively, in hydrolyzing oNp-gal-6-P than wild-type CelB.

DISCUSSION

CelB Structural Model and Active Site. A structural model of *P. furiosus* β-glucosidase with a resolution of 3.3 Å has been constructed based on X-ray diffraction of its crystals, followed by molecular replacement of the 3D structures of *L. lactis* LacG and *S. solfataricus* LacS. Although this resolution did not allow a detailed analysis of the entire protein, the structure of the active site was analyzed in more detail, since it could be established with a high accuracy (Figures 1 and 2) and confirms earlier predictions that CelB is a retaining enzyme (*22*, *24*, *25*).

The central fold of CelB is a $(\beta\alpha)_8$-barrel (Figure 1), first found in chicken triosephosphate isomerase (*36*). This fold has been predicted for all family 1 glycosyl hydrolases and has indeed been observed for all established structures of family 1 glycosyl hydrolases (*6*−*8*, *10*, *11*). A tetrameric subunit organization has been observed for all hyperthermostable representatives of family 1 β-glycosidases, including *Thermosphaera aggregans* and *Sulfolobus solfataricus* as well (*11*, *12*). This oligomerization may contribute to stability, since mesophilic and thermophilic family 1 enzymes are mainly active as monomers or dimers (*6*, *8*). However, higher degrees of oligomerization have been observed (*9*). The positions of the active site residues in CelB, LacG, and LacS are highly conserved (Figure 2). The glutamate residues are located at the C-terminal ends of the fourth and seventh β-strands of the barrel and identify the enzymes as a member of the glycosyl hydrolase 4/7-superfamily (*13*). The opening at the entrance of the substrate channel toward the active site, which results from a bend in the helix of the fifth βα-unit, is similar to the one found in *S. solfataricus* LacS. However, the unusual amino acid sequence, WWFF, at the corresponding positions (287−290) in *S. solfataricus* LacS is not present in CelB.

The family 1 6-P-β-galactosidases have evolved a binding site for the phosphate group of galactose-6-phosphate. In *L. lactis* LacG, Ser428, Lys435, and Tyr437 were found to interact with the phosphate group (*10*). The corresponding residues in CelB (Glu417, Met424, Phe426) are highly conserved in the β-glycosidases (Figure 3). The negatively charged Glu417 would have a repulsive action on the phosphate group of an incoming 6-phosphoglycoside. The mutations E417S, M424K, and F426Y were therefore introduced to create a phosphate binding site in CelB that may affect the kinetics of hydrolysis of 6-phosphoglycosides.

Characterization of Mutants. The structure of the active site confirms that CelB hydrolyzes its substrates by a retaining mechanism in which a covalent intermediate is formed (*24*). In the catalysis cycle, the C1 atom of an incoming glycoside is attacked by the nucleophilic glutamate (Glu372 in CelB). The glycosidic bond is broken, and a covalent enzyme−substrate intermediate is formed. An incoming water molecule hydrolyzes the covalent bond, and the enzyme returns to the initial state (*37*). In wild-type CelB, the hydrolysis of the glycosyl-enzyme intermediate is the

rate-limiting step in the reaction with oNp-glu. The rate is dependent on the ionization state of the general acid/base catalytic residue that abstracts a proton of an incoming water, which, in turn, attacks the covalent bond between the enzyme and the sugar moiety (*3*).

The substitutions M424K and F426Y had relatively small effects on the hydrolysis of oNp-glu and oNp-gal. The pH profile of F426Y was identical to wild-type CelB, as was expected since this substitution is charge-neutral. The substitution M424K did introduce a positive charge into the active site, and in M424K and the double mutant M424K/F426Y, the pH optimum had shifted to a higher value (Figure 4). Probably, the introduced positive charge at this position results in an increase of the pK_a value of the nucleophilic glutamate, as can be deduced from the shape of the curve at low pH values (*35*). The distance between Glu372 and Lys424 is too large (~15 Å) for a direct interaction, but there could be an influence of the charge or size on the local conformation of the enzyme. Residue Lys424 is also present at the corresponding position in the wild-type β-glucosidases from *Thermotoga maritima* and *Caldocellum saccharolyticum*. Remarkably, the pH optimum of CelB M424K was similar to the one determined for the β-glucosidase of *T. maritima* (*38*). In general, the substitutions M424K and F426Y caused a decrease in catalytic efficiencies on oNp-glu and oNp-gal, except for F426Y on oNp-gal. For both substrates, the double mutant showed the combined effects of the single mutants, an indication for the proper folding of the mutant's active sites (Table 2).

Substitution E417S influenced the catalytic properties of CelB significantly. Although wild-type CelB shows the fastest hydrolysis of oNp-glu and oNp-gal at pH 5.0, it retains 90% of its activity between pH 4.5 and 6.0. E417S and the triple mutant showed narrow curves with sharp optima at pH 4.5 and 5.0, respectively (Figure 4). It is likely that Glu417, like Lys424, is too far away from the catalytic glutamates Glu207 and Glu372 for a direct interaction.

The glutamate residue in *Bacillus polymyxa* β-glucosidase BglA that corresponds to Glu417 in CelB was found to interact with the hydroxyl group of C4 and C6 of a gluconate molecule bound in the active site. Since glucose and galactose are identical molecules except for the position of the hydroxyl group on the C4 position, equatorial in glucose and axial in galactose, it is believed that the position of this glutamate residue is partly responsible for the difference in substrate specificity between a glucosidase and a galactosidase (*9*). Furthermore, Glu417 is important for the stabilization of the glycosyl-enzyme intermediate in the hydrolysis cycle (*39*). The effects of the disturbed interaction with the C4 hydroxyl group become clear from the kinetic parameters on oNp-glu and oNp-gal in E417S and the triple mutant. The K_M values for oNp-glu and oNp-gal were found to increase 100-fold and 20-fold, respectively. In *S. solfataricus* LacS, Glu432, corresponding to residue Glu417 in CelB, was postulated to interfere with the carboxylate group of substrates such as galuronic acid or glucuronic acid (*11*). However, the replacement of Glu417 by a serine residue did not enable CelB to hydrolyze *p*-nitrophenol-β-D-glucuronate (data not shown).

The substitutions M424K and F426Y did not improve the hydrolysis of oNp-gal-6-P. M424K increased the optimum pH for catalysis to pH 4.5, but, like F426Y, did not result in a higher catalytic efficiency. Substitution E417S has the largest effect on the hydrolysis of phosphorylated galactosides. On one hand, this substitution enables the hydrolysis of oNp-gal-6-P at a higher pH value compared to oNp-glu or oNp-gal. On the other hand, it lowers the Michaelis constant for the phosphorylated substrate and increases the k_{cat}. As a consequence, mutant E417S and the triple mutant have 4.9 and 2.7 times improved efficiencies for the hydrolysis of phosphorylated galactose, respectively. The combination of the three mutations in CelB triple results in the lowest K_m, and the highest pH optimum. Apparently, the lysine contributes to the pH optimum for the hydrolysis of phosphorylated galactose, as can be concluded from the difference in the pH optima of CelB E417S and the CelB triple mutant with oNp-gal-6-P as a substrate.

When compared to the 6-phospho-β-galactosidase LacG, there is a striking difference in K_M values (Table 3). At 90 °C, the Michaelis constant K_M of the triple mutant is 40-fold higher than the K_M of LacG for oNp-gal-6P at 25 °C. This could be due to a difference in temperature, although this is not predictable, as has been found in other studies (*40*). In addition to the substitution of residues in the active site, additional replacements of second shell residues are probably needed to further optimize CelB for the hydrolysis of phosphorylated galactosides. However, it can be concluded that the residue at position 417 in CelB has a major determining role in the specificity of family 1 glycosyl hydrolases. A glutamate at this position will make the enzyme efficient in the hydrolysis of β-glycosides, while a serine at this position is mainly responsible for 6-phospho-β-galactosidase activity.

Based on a comparison of the 3D structure of *L. lactis* 6-phospho-β-galactosidase LacG and the model of the *P. furiosus* β-glucosidase, three substitutions were introduced in the active site of CelB, with the aim to increase its activity toward phosphorylated galactose. The roles of the three residues in hydrolysis have been evaluated and were found to be of importance for the enzyme's substrate affinity, pH-optimum, and hydrolysis rate. In conclusion, a molecular basis has been provided for the fundamental difference in substrate specificity between members of family 1 glycosyl hydrolases that hydrolyze specifically phosphorylated and nonphosphorylated substrates. Moreover, the CelB structural model has been verified and may provide a basis for further understanding of catalysis and stability at extreme high temperatures.

ACKNOWLEDGMENT

We thank Wilfried Voorhorst and Jurre Koning for their contributions in early stages of this work, Willem van Berkel for valuable discussions, and Laurence Pearl and Mose Rossi for sharing the *S. solfataricus* LacS dataset prior to publication.

REFERENCES

1. Davies, G., and Henrissat, B. (1995) *Structure 3*, 853−859.
2. Henrissat, B. (1999) http://afmb.cnrs-mrs.fr/~pedro/CAZY/ghf.html.
3. Kempton, J. B., and Withers, S. G. (1992) *Biochemistry 31*, 9961−9969.
4. Sinnott, M. L. (1990) *Chem. Rev. 90*, 1171−1202.

5. Wang, Q., Trimbur, D., Graham, R., Warren, R. A., and Withers, S. G. (1995) *Biochemistry 34*, 14554−14562.
6. Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J., and Hughes, M. A. (1995) *Structure 3*, 951−960.
7. Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S., and Henrissat, B. (1997) *Structure 5*, 663−675.
8. Wiesmann, C., Beste, G., Hengstenberg, W., and Schulz, G. E. (1995) *Structure 3*, 961−968.
9. Sanz Aparicio, J., Hermoso, J. A., Martinez Ripoll, M., Lequerica, J. L., and Polaina, J. (1998) *J. Mol. Biol. 275*, 491−502.
10. Wiesmann, C., Hengstenberg, W., and Schulz, G. E. (1997) *J. Mol. Biol. 269*, 851−860.
11. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., and Pearl, L. H. (1997) *J. Mol. Biol. 271*, 789−802.
12. Chi, Y. I., Martinez-Cruz, L. A., Jancarik, J., Swanson, R. V., Robertson, D. E., and Kim, S. H. (1999) *FEBS Lett. 445*, 375−383.
13. Jenkins, J., Lo Leggio, L., Harris, G., and Pickersgill, R. (1995) *FEBS Lett. 362*, 281−285.
14. Street, I. P., Kempton, J. B., and Withers, S. G. (1992) *Biochemistry 31*, 9970−9978.
15. Trimbur, D. E., Warren, R. A., and Withers, S. G. (1992) *J. Biol. Chem. 267*, 10248−10251.
16. Voorhorst, W. G. B., Eggen, R. I. L., Luesink, E. J., and De Vos, W. M. (1995).
17. Moracci, M., Capalbo, L., Ciaramella, M., and Rossi, M. (1996) *Protein Eng. 9*, 1191−1195.
18. Febbraio, F., Barone, R., D'Auria, S., Rossi, M., Nucci, R., Piccialli, G., De Napoli, L., Orru, S., and Pucci, P. (1997) *Biochemistry 36*, 3068−3075.
19. Kengen, S. W., Luesink, E. J., Stams, A. J., and Zehnder, A. J. (1993) *Eur. J. Biochem. 213*, 305−312.
20. Lebbink, J. H. G., Kaper, T., Kengen, S. W. M., Van der Oost, J., and De Vos, W. M. (2000) *Methods Enzymol*. (in press).
21. Lebbink, J. H. G. (1999) Ph.D. Thesis, p 192, Wageningen University, Wageningen, The Netherlands.
22. Voorhorst, W. G., Eggen, R. I., Luesink, E. J., and de Vos, W. M. (1995) *J. Bacteriol. 177*, 7105−7111.
23. Bauer, M. W., Bylina, E. J., Swanson, R. V., and Kelly, R. M. (1996) *J. Biol. Chem. 271*, 23749−23755.
24. Bauer, M. W., and Kelly, R. M. (1998) *Biochemistry 37*, 17170−17178.
25. Fischer, L., Bromann, R., Kengen, S. W., de Vos, W. M., and Wagner, F. (1996) *Bio/Technology 14*, 88−91.
26. Boon, M. A., Van der Oost, J., De Vos, W. M., Jansen, A. E. M., and Van't Riet, K. (1998) *Appl. Biochem. Biotechnol. 75*, 269−278.
27. Petzelbauer, I., Nidetzky, B., Haltrich, D., and Kulbe, K. D. (1999) *Biotechnol. Bioeng. 64*, 322−332.
28. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) (Nolan, C., Ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res. 16*, 7351−7367.
30. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
31. Jancarik, J., Scott, W. G., Milligan, D. L., Koshland, D. E., Jr., and Kim, S. H. (1991) *J. Mol. Biol. 221*, 31−34.
32. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
33. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science 235*, 458−460.
34. McCarter, J. D., and Withers, S. G. (1994) *Curr. Opin. Struct. Biol. 4*, 885−892.
35. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) *Biochemistry 35*, 9958−9966.
36. Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D., and Waley, S. G. (1975) *Nature 255*, 609−614.
37. White, A., and Rose, D. R. (1997) *Curr. Opin. Struct. Biol. 7*, 645−651.
38. Gabelsberger, J., Liebl, W., and Schleifer, K. (1993) *Appl. Microbiol. Biotechnol. 40*, 44−52.
39. Namchuk, M. N., and Withers, S. G. (1995) *Biochemistry 34*, 16194−16202.
40. Lebbink, J. H. G., Kaper, T., Bron, P., Van der Oost, J., and Vos, W. M. (2000) *Biochemistry 39*, 3656−3665.
41. Hengstenberg, W., Kohlbrecher, D., Witt, E., Kruse, R., Christiansen, I., Peters, D., Pogge von Strandman, R., Stadtler, P., Koch, B., and Kalbitzer, H. R. (1993) *FEMS Microbiol. Rev. 12*, 149−164.
42. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.

BI992463R