# II. ARTICLES

# II.1. ARTICLE PRIMER

**Clonatge, expressió i cristal·lització de la β-glucosidasa Bgl3.**

## II.1.a. CLONATGE, EXPRESSIÓ I CRISTAL·LITZACIÓ DE LA β-GLUCOSIDASA BGL3

### "Cloning, overexpression, crystallization and preliminary X-ray analysis of a family 1 β–glucosidase from *Streptomyces*."

Aquest treball descriu el clonatge i la sobreexpressió en *Escherichia coli* d'una β-glucosidasa (Bgl3) intracel·lular de *Streptomyces* sp. amb l'objectiu de disposar un sistema eficient i reproduïble d'obtenció de proteïna per endegar els estudis de cristal·lització que van portar finalment a resoldre la seva estructura tridimensional per difracció de raigs X. Després de provar diferents estratègies/vectors, els millors resultats en termes d'estabilitat dels clons, expressió en forma soluble i rendiment de proteïna es van obtenir amb el sistema pET (Novagen). La introducció d'una cua d'histidines a l'extrem N-terminal de la proteïna ha possibilitat la seva purificació a homogeneïtat en una sola etapa cromatogràfica, amb rendiments de 150-200 mg de proteïna pura per litre de cultiu. La proteïna expressada de forma recombinant presentava les mateixes característiques i propietats de l'enzim natiu: és una glicosidasa (52,6 kDa) que actua amb retenció de configuració, pot hidrolitzar diversos disacàrids i oligosacàrids i dur a terme reaccions de transglicosidació (activitat sintètica). Inicialment, els intents de cristal·lització de la proteïna expressada en *E. coli* van rendir, en el millor dels casos, cristalls petits i de baixa qualitat de difracció. La proteòlisi limitada amb tripsina va permetre millorar tant el creixement dels cristalls com la seva estabilitat. L'anàlisi per SDS-PAGE de la proteïna tripsinitzada va posar de manifest que s'obtenien com a productes límit dos fragments de 35 i 20 kDa aproximadament. Una determinació més acurada del punt de tall de la tripsina es va determinar combinant la seqüenciació N-terminal amb l'espectrometria de masses per MALDI-TOF. Cristalls de Bgl3 tripsinitzada van créixer en una solució de sulfat amònic en tampó HEPES (pH 7,5) a 20 ºC mitjançant el mètode de la difusió de vapor en gota penjant. L'anàlisi preliminar dels cristalls va proporcionar les següents dades: (1) pertanyen al grup espaial ortoròmbic I222; (2) les dimensions de la cel·la unitat a temperatura ambient són a=101,6, b=113,4 i c=187,5 Å; i (3) contenen dues molècules de proteïna per unitat asimètrica. Aquests cristalls difractaven més enllà de 2,7 Å de resolució utilitzant una font de raigs X convencional d'ànode giratori però no eren estables en el sentit que patien un canvi significatiu en les dimensions de cel·la durant la recollida de dades. La millora de l'estabilitat dels cristalls i de la qualitat de la seva difracció es va aconseguir utilitzant la radiació de sincrotó sobre cristalls "congelats" (-170 ºC) en una solució de sacarosa al 30%. Els cristalls congelats van difractar més enllà de 1,69 Å i es va poder recollir un conjunt de dades suficientment complet (97,7%) per a la determinació de l'estructura mitjançant reemplaçament molecular (utilitzant la glucosidasa de *Trifolium repens* -codi PDB, 1CBG- com a model).

Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Cloning, overexpression, crystallization and preliminary X-ray analysis of a family 1 β-glucosidase from *Streptomyces*

**Alicia Guasch,[a]* Miquel Vallmitjana,[b] Rosa Pérez,[a] Enrique Querol,[b] Josep A. Pérez-Pons[b] and Miquel Coll[a]**

[a]Institut de Biologia Molecular CSIC Jordi Girona 18-26, 08034 Barcelona, Spain, and [b]Departament de Bioquímica i Biologia Molecular and Institut de Biologia Fonamental, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

Correspondence e-mail: agmcri@cid.csic.es

Received 30 September 1998
Accepted 22 October 1998

An intracellular β-glucosidase (Bgl3) from *Streptomyces* sp. has been cloned and overexpressed in *Escherichia coli*. The introduction of a His tag at the N-terminal end of the protein has allowed its purification to homogeneity by a single chromatographic step, with yields of 150–200 mg of pure protein per litre of *E. coli* culture. The enzyme (52.6 kDa) is a retaining glycosidase able to hydrolyze a wide range of disaccharides and oligosaccharides and to perform transglycosylation. Crystals of recombinant Bgl3 have been grown from an ammonium sulfate solution using the hanging-drop vapour-diffusion method at 293 K. The crystals belong to the orthorhombic space group *I*222 with unit-cell dimensions $a$ = 101.6, $b$ = 113.4 and $c$ = 187.5 Å at room temperature and contain two molecules per asymmetric unit. A full 1.69 Å resolution diffraction data set (97.7% completeness) has been collected from frozen crystals in a solution containing 30% sucrose, using synchrotron radiation.

## 1. Introduction

β-Glucosidases (β-glucoside glucohydrolase, E.C. 3.2.1.21) are enzymes which act on carbohydrates and catalyze the transfer of glucosyl groups between oxygen nucleophiles by means of a general acid mechanism with retention of the anomeric configuration and through the formation of a covalent glycosyl–enzyme intermediate (McCarter & Withers, 1994; White & Rose, 1997). Under physiological conditions, such a transfer reaction generally results in the hydrolysis of a β-glucosidic bond from disaccharides, oligosaccharides or conjugated glucosides. Thus, β-glucosidases are widely distributed in nature, playing different roles. Their involvement in enzymatic systems is relevant to plant biomass degradation and molecular pathologies such as Gaucher's disease in humans. Another important feature of β-glucosidases is their transglycosylation activity (*i.e.* transfer of glycosidic residues), which could provide a valuable tool for the specific enzymatic synthesis of sugars and sugar derivatives. On the basis of sequence similarity and hydrophobic cluster analysis, β-glucosidases have been classified in families 1 and 3 of the large group of glycosyl hydrolases, which includes more than 60 families (Henrissat & Davies, 1997). According to this classification, family 1 includes 6-phospho-β-galactosidases (E.C. 3.2.1.85), 6-phospho-β-glucosidases (E.C. 3.2.1.86), lactase-phlorizin hydrolases (E.C. 3.2.1.108) and myrosinases (E.C. 3.2.3.1) in addition to β-glucosidases. With regard to the overall folding, family 1 enzymes are included into the clan or superfamily GH-A, characterized by an $(\alpha/\beta)_8$ barrel structure, whereas no crystallographic data are available for any enzymes belonging to family 3. The GH-A clan contains at least 11 glycosyl hydrolase families representing a wide variety of enzymatic activities which include, among others, endoglucanases, lichenases, glucosylceramidases and xylosidases (Bairoch, 1998). Such a picture means that the α/β barrel fold is a tertiary structure able to accommodate a diversity of both amino-acid sequences and substrate specificities. Hence, work is still necessary to fully understand the elements governing structure–function relationships in this superfamily of glycosyl hydrolases. In recent years, five crystal structures of enzymes belonging to family 1 glycosyl hydrolases have been reported: a cyanogenic β-glucosidase from *Trifolium repens* (Barrett *et al.*, 1995), a phospho-β-galactosidase from *Lactococcus lactis* (Wiesmann *et al.*, 1995), a myrosinase from *Sinapis alba* (Burmeister *et al.*, 1997), a thermostable β-galactosidase from *Sulfolobus solfataricus* (Aguilar *et al.*, 1997) and a β-glucosidase from *Bacillus polymyxa* (Sanz-Aparicio *et al.*, 1998).

The gene encoding Bgl3 glucosidase from *Streptomyces* sp. QM-B814 was initially cloned by functional complementation of a β-glucosidase-negative mutant of *Streptomyces lividans* and the enzyme, as expressed in this host system, was characterized biochemically. It was shown to be a retaining glycosidase with an exo-like action pattern, releasing glucose units from the non-reducing end of diverse disaccharides and oligosaccharides (Pérez-Pons *et al.*, 1994). Upon sequence-similarity analysis, it

has been classified into the family 1 glycosyl hydrolases, showing identity indexes from 35 to 55% with respect to different members of that family. Protein crystallization trials were initially carried out with protein purified from crude cell extracts of recombinant *S. lividans* by six chromatographic steps with an overall yield of 10% (unpublished results). Large-scale fermenter cultures were necessary to obtain amounts of pure protein sufficient for extensive crystallization screening. Crystals were obtained from several conditions by the vapour-diffusion method but all of them exhibited very poor X-ray diffraction properties, probably owing to their small size. In order to overcome the time-consuming and low-yielding protein preparation process from *Streptomyces*, attempts to express Bgl3 in *E. coli* were undertaken, including extracellular expression by fusion to several signal peptides from *E. coli* and *Bacillus licheniformis*.

In this communication, we report the crystallization and preliminary X-ray analysis of the Bgl3 glucosidase of *Streptomyces* sp. QM-B814 as expressed in *E. coli* using the pET system (Novagen), which allowed very high yields of enzyme after a one-step purification scheme. This work constitutes the first step towards a detailed three-dimensional structure analysis of this enzyme at high resolution, which will afford new insights into the overall architecture of the GH-A clan and the structural elements determining the fine tuning of substrate specificity in the large glycosyl hydrolase family 1.

## 2. Experimental

### 2.1. Cloning, expression and purification of Bgl3 from *E. coli*

The $\beta$-glucosidase gene from *Streptomyces* sp. QM-B814, including its downstream region which may function as a transcription terminator since it contains two inverted repeats (CGGTGCGGCAC and GGCCCGCCCCCG starting at 8 and 37 nucleotides from the stop codon, respectively), was cloned into the T7 polymerase expression vector pET-21d(+) (Novagen) and overexpressed in *E. coli* strain BL21(DE3) (Novagen). Likewise, a His-tag coding sequence was fused to the start codon (GTG) of the Bgl3 reading frame obtaining the recombinant plasmid pET21-HBG3, which yields a protein with an extended N-terminal formed by the sequence MHHHHHHGIH and a deduced molecular mass of 53.6 kDa.

In order to obtain reproducible yields of protein, each production batch was started by transforming competent *E. coli* BL21(DE3) cells with the plasmid pET21-HBG3. One transformant colony grown in LB solid medium containing ampicillin (100 $\mu$g ml$^{-1}$), was picked, inoculated into 3 ml of 2xYT broth containing ampicillin (200 $\mu$g ml$^{-1}$) (Sambrook *et al.*, 1989) and shaken overnight at 310 K to an $OD_{595}$ of about 0.6; aliquots of this culture were used to inoculate larger volumes of the same growth medium (usually 200 ml) that were shaken (300 rev min$^{-1}$) for 8–10 h at 310 K. Following such a procedure, expression of recombinant Bgl3 was achieved constitutively without addition of isopropyl-$\beta$-thiogalactopyranoside (IPTG). The cells were harvested and disrupted essentially according to the protocol described in the pET System Manual (Novagen); the sonication step was omitted (see §3). For protein purification, cell-free extracts were centrifuged (39000*g*, 30 min at 277 K) and the supernatant (soluble fraction) applied to a 5 ml Hi-Trap Chelating Sepharose (Pharmacia) column previously charged with $Ni^{2+}$. Chromatography was carried out according to the manufacturer's specifications and Bgl3-containing fractions were pooled, concentrated and desalted by ultrafiltration at 277 K using 50 m*M* sodium phosphate (pH 7.0) as desalting buffer. Using this procedure a homogeneous (>99%) protein sample was obtained, as judged by sodium dodecyl sulfate–polyacrylamide-gel electrophoresis (SDS–PAGE) and Coomassie Brilliant Blue R-250 staining. The protein concentration was determined by the dye-binding method of Bradford (1976) using BSA as a standard. The pure enzyme was kept at 277 K for short periods or stored at 253 K in the presence of 45% glycerol for prolonged periods.

### 2.2. Crystallization

Initial crystallization conditions were obtained by the hanging-drop vapour-diffusion method (McPherson, 1982) with a broad screening according to the conditions described by Jancarik & Kim (1991) and Cudney *et al.* (1994) included in Crystal Screening Kits I and II (Hampton Research, California, USA). The droplets were prepared by mixing 2 $\mu$l of protein solution (12 mg ml$^{-1}$) with an equal volume of mother liquor. The trials were carried out at a constant temperature of either 277 or 293 K. Another set of trials was prepared using a trypsinized Bgl3 sample (see §3). Time-course trypsin (Boehringer) digestion was performed at a protease:Bgl3 ratio of 1:100 in 45 m*M* Tris–HCl buffer (pH 8.1) at 310 K. Reactions were stopped by adding 1 m*M* Pefabloc SC (Boehringer) and analyzed by means of SDS–PAGE and Coomassie Blue staining. The trypsin cleavage site was determined by microsequencing the first ten N-terminal amino-acid residues from each digestion product using the Edman degradation method combined with mass-spectrometry analysis (MALDI–TOF system).

### 2.3. X-ray crystallography

Diffraction data were collected both at 293 K using an in-house rotating-anode X-ray source with $\lambda = 1.5418$ Å and at 103 K on the beamline BW7A of the EMBL outstation at the Deutches Elektronensynchrotron (Hamburg, Germany), $\lambda = 0.99$ Å. In both cases a MAR Research 300 mm imaging plate was used as detector. Determination of unit-cell parameters, space group, integration of reflection intensities and data scaling were performed using *DENZO* and *SCALEPACK* (Otwinowski, 1993).

## 3. Results and discussion

The crystallization of the Bgl3 glucosidase has proved to be more difficult than expected, probably owing to intrinsic structural features of the protein and not to the presence of impurities or sample heterogeneity. Nevertheless, it deserves to be mentioned that the inclusion of a sonication step during the production of the crude extract seems to interfere negatively with the crystal growth, probably as a consequence of the solubilization of misfolded Bgl3 protein leading to sample microheterogeneity. In any case, the cloning and overexpression of Bgl3 in *E. coli* using the pET vector and the fusion of a His tag at the N-terminus allowed improvement in both the availability of large amounts of protein and the simplicity of its purification. In fact, pure protein yields increased from about 70 mg per 15 l fermenter culture of *S. lividans* (and six chromatographic steps) to 150–200 mg l$^{-1}$ of *E. coli* culture following the one-step purification procedure described in §2. In contrast, the fusion of the His tag to the C-terminal end of Bgl3 glucosidase was deleterious for its expression, as was any attempt to obtain secretion by fusing different signal peptides to the N-terminal end.

As stated in §1, at the beginning of this work we obtained small poorly diffracting

crystals of native Bgl3 purified from *S. lividans*. Attempts were made to improve the crystal size and diffraction quality by seeding, changes in buffer, the use of additives such as alcohols, organic solvents and detergents, and the presence of substrates or inhibitors. None of these trials led to better and useful results. Likewise, many attempts to obtain crystals of *E. coli*-expressed native Bgl3 were carried out, but all were unsuccessful. It is known that in some cases limited proteolysis improves both crystal growth and stability (Chitarra *et al.*, 1995). Thus, protease digestions were checked by using different enzymes, obtaining the best and reproducible results with trypsin reactions. After 1 h digestion two bands of about 35 and 20 kDa were observed, and no further proteolysis was detected upon extending the reaction for up to 6 h. Following the methodology described in §2, we determined that the major trypsin-cleavage sites correspond to Ser315 and Arg330, yielding two fragments of 33560 and 16608 Da consistent with the mass estimated by SDS–PAGE analysis. Furthermore, trypsinized samples were loaded onto a Sephacryl S-100 gel-filtration column previously equilibrated with 0.2 *M* ammonium acetate buffer (pH 6.0) and only one peak eluted from the column. This peak was active using *p*-nitrophenyl-β-D-glucopyranoside as a substrate, indicating that trypsin digestion gives a stable and functional Bgl3 core enzyme. Upon sequence alignment between Bgl3 and family 1 glycosidases whose structures are already known (see §1), it can be inferred that the trypsin-cleavage sites are situated within a region forming a loop. In addition, such a loop would be larger in the *Streptomyces* enzyme.

**Table 1**
Data collection and statistics.

| Temperature (K) | 293 | 103 |
|---|---|---|
| Unit-cell dimensions (Å) | $a$ = 101.6, $b$ = 113.4, $c$ = 187.5 | $a$ = 96.7, $b$ = 111.3, $c$ = 185.3 |
| Resolution range (Å) | 40–2.7 | 40–1.69 |
| Number of unique reflections | 26565 | 108818 |
| Completeness (%) | | |
| All data | 88.2 | 97.7 |
| Highest resolution shell† | 84.3 | 68.4 |
| Average redundancy | 2.3 | 4.4 |
| $\langle I/\sigma(I)\rangle$ | | |
| All data | 23.4 | 34.3 |
| Highest resolution shell† | 22.7 | 5.87 |
| $R_{\rm merge}$‡ (%) | | |
| All data | 9.7 | 3.8 |
| Highest resolution shell† | 11.3 | 23.2 |

† Highest resolution shell was 2.8–2.7 and 1.72–1.69 Å for 293 and 103 K crystals, respectively. ‡ $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

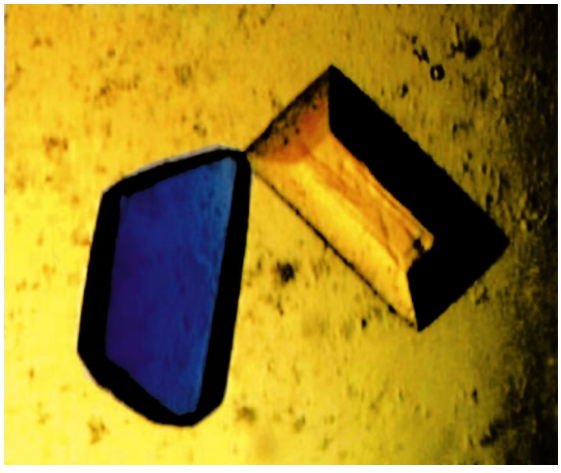


**Figure 1**
Crystals of trypsinized β-glucosidase, grown in 1.9 *M* ammonium sulfate, 0.1 *M* HEPES (pH 7.5) at 293 K; dimensions are approximately 0.4 × 0.4 × 0.2 mm.

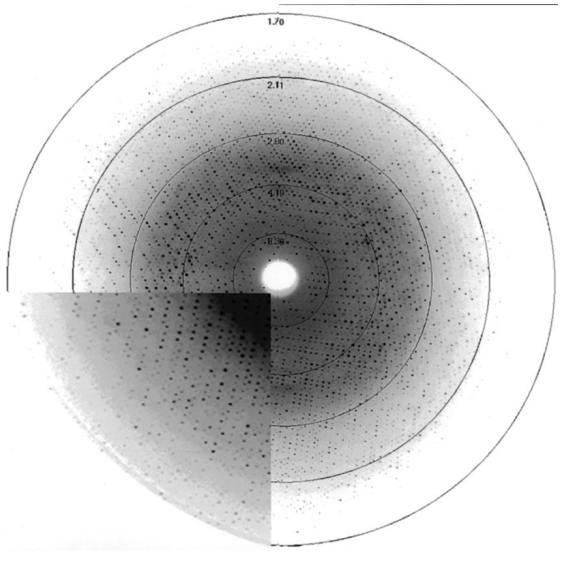



**Figure 2**
1° rotation image obtained from a frozen crystal using synchrotron radiation. Resolution rings are labelled in Å.

Crystals were obtained in 1.9 *M* ammonium sulfate, 0.1 *M* HEPES (pH 7.5) at 293 K. Single well shaped colourless bipyramidal crystals appeared in about three weeks (Fig. 1), reaching a full size of about 0.4 × 0.4 × 0.2 mm after four months. The crystals diffracted beyond 2.7 Å resolution using a rotating-anode source at 293 K. A data set was collected, but a significant change of the cell dimensions (5 Å in $a$) was observed after 50 frames. The indexing procedure carried out with *DENZO* (Otwinowski, 1993) indicated that the crystals belong to space group $I222$, with unit-cell dimensions of $a$ = 101.6, $b$ = 113.4 and $c$ = 187.5 Å. Flash-freezing after crystal transfer to a solution containing 30% sucrose and 2.25 *M* ammonium sulfate (pH 7.5) improved crystal stability and diffraction quality (Fig. 2). Using synchrotron radiation, frozen crystals diffracted beyond 1.69 Å resolution and the unit-cell dimensions were $a$ = 96.7, $b$ = 111.3 Å and $c$ = 185.3 Å. Table 1 shows the data-collection statistics. The unit-cell volume is consistent with two monomers in the asymmetric unit, yielding an acceptable specific volume ($V_m$) of 2.48 Å$^3$ Da$^{-1}$ and a corresponding solvent content of 40% (Matthews, 1968). The structure determination by molecular replacement using *Trifolium repens* cyanogenic β-glucosidase (PDB entry 1CBG; Barrett *et al.*, 1995) is in progress.

The *E. coli* expression system has also proven to be useful for the overproduction of Bgl3 point mutants obtained by site-directed mutagenesis, and several different mutants have been produced. The crystallization and structure determination of these mutant enzyme forms will provide useful tools in the evaluation and understanding of the principles that determine substrate specificity. With this aim, we have also prepared a series of co-crystals with non-hydrolyzable substrate analogues and heavy-atom containing inhibitors.

This work was supported by the Ministerio de Educación y Ciencia (grants PB95-0224 to MC and BIO 97-0511-CO2-01 to EQ) and the Generalitat de Catalunya [Centre de Referència en Biotecnologia

(CERBA) and grant 1997SGR-275 to MC]. Data collection at EMBL-DESY was supported by the EU Large Installations Project CHGECT-93-0040. We thank F. Canals for his assistance in performing mass spectrometry and amino-acid sequencing analyses. MV is the recipient of a predoctoral scholarship from CERBA.

## References

Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). *J. Mol. Biol.* **271**, 789–802.

Bairoch, A. (1998). *SWISS-PROT Protein Sequence Data Bank*, http://expasy.hcuge.ch/cgi-bin/lists?glycosid.txt.

Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. & Hughes, M. A. (1995). *Structure*, **3**, 951–960.

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.

Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. (1997). *Structure*, **5**, 663–675.

Chitarra, V., Souchon, H., Spinelli, S., Juy, M., Beguin, P. & Alzari, P. M. (1995). *J. Mol. Biol.* **248**, 225–232.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.

Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.

Jancarik, J. & Kim, S. H. (1991). *J. Appl. Cryst.* **24**, 409–411.

McCarter, J. D. & Withers, S. G. (1994). *Curr. Opin. Struct. Biol.* **4**, 885–892.

McPherson, A. (1982). *The Preparation and Analysis of Protein Crystals*. New York: John Wiley.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Pérez-Pons, J. A., Cayetano, A., Rebordosa, X., Lloberas, J., Guasch, A. & Querol, E. (1994). *Eur. J. Biochem.* **223**, 557–565.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Sanz-Aparicio, J., Hermoso, J. A., Martínez-Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). *J. Mol. Biol.* **275**, 491–502.

White, A. & Rose, D. R. (1997). *Curr. Opin. Struct. Biol.* **7**, 645–651.

Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). *Structure*, **3**, 961–968.

# II.2. ARTICLE SEGON

**Estudis cinètics i mecanisme catalític de la β-glucosidasa Bgl3.**

## II.2.a. ESTUDIS CINÈTICS I MECANISME CATALÍTIC DE LA β-GLUCOSIDASA Bgl3

### "Mechanism of the Family 1 β-Glucosidase from *Streptomyces* sp: Catalytic Residues and Kinetic Studies."

Per anàlisis de similitud de seqüència la Bgl3 s'ha inclòs dins de la família 1 de glicosilhidrolases. Aquest article explica el treball realitzat per a identificar els residus catalítics de la β–glucosidasa Bgl3 d'*Streptomyces* sp. i caracteritzar acuradament les propietats cinètiques de l'enzim salvatge amb la finalitat, comparant-les amb altres β-glucosidases de la família 1, d'aprofundir en l'estudi dels detalls mecanístics que governen l'activitat enzimàtica. Aquests enzims catalitzen la transferència de grups glucosil entre oxígens nucleofílics a través d'un mecanisme de doble desplaçament amb formació d'un intermediari covalent glicosil-enzim. Per aquest mecanisme, generalment es requereixen en el centre actiu dos residus carboxilat: un d'ells actuaria com a nucleòfil, atacant el centre anomèric del substrat (corresponent a l'extrem no reductor) i desplaçant el grup sortint per a formar l'intermediari covalent, i l'altre com a catalitzador àcid/base, proporcionant l'assistència per alliberar el grup sortint (protonació) i per l'atac posterior d'una molècula d'aigua sobre l'intermediari covalent (desprotonació) que dóna, de nou, enzim lliure i un producte amb retenció de la configuració anomèrica.

Els alineaments de seqüència i el modelat molecular van permetre proposar els residus Glutàmic-383 i Glutàmic-178 com a candidats a nucleòfil i catalitzador àcid/base, respectivament. Aquests residus van ser substituïts per glutamina (isostèric amb pèrdua de càrrega) i per alanina (pèrdua de cadena lateral i de càrrega) mitjançant tècniques de mutagènesi dirigida. Els enzims mutants es van expressar en el sistema *E. coli*-pET (Novagen) i es van purificar per cromatografia d'afinitat per $Ni^{2+}$ gràcies a la fusió d'una cua d'histidines a l'extrem N-terminal de les proteïnes. El correcte plegament i l'estabilitat de les formes mutants obtingudes es va avaluar seguint per fluorescència la desnaturalització amb urea (0-8 M). Així, els canvis en les propietats catalítiques es van poder atribuir a la substitució dels residus escollits. La substitució del Glu-178 per Gln i Ala va donar enzims amb una eficiència catalítica reduïda 250 i 3500 vegades, respectivament; la mutació del Glu-383 va donar enzims encara més inactius (reducció de $10^5$-$10^6$ vegades). El paper funcional d'ambdós residus es va confirmar per la metodologia del rescat químic, basada en l'activació dels mutants Ala inactius mitjançant azida com a nucleòfil exogen. Els productes resultants de la reactivació es van analitzar per $^1$H-RMN. El mutant E178A va rendir β–glucosil azida amb un increment de 200 vegades en el valor de $k_{cat}$ pel 2,4-dinitrofenil glucòsid; el mutant E383A amb el mateix substrat va donar α-glucosil azida com a producte i un increment de 100 vegades en la $k_{cat}$ a 1 M azida. Els resultats confirmen inequívocament que el residu Glu-178 és el catalitzador àcid/base (configuració β de l'adducte) i el Glu-383 el nucleòfil catalític (configuració α de l'adducte).

L'especificitat de substrat es va estudiar realitzant cinètiques d'estat estacionari amb diferents p-nitrofenil-β-D-glicòsids. Els resultats van posar de manifest que la major eficiència catalítica ($k_{cat}/K_M$) de l'enzim salvatge es donava amb els derivats fucòsid (sense inhibició per excés de substrat) i glucòsid (amb forta inhibició per substrat); d'altra banda, l'activitat glucosidasa és 10 vegades superior a l'activitat galactosidasa. Per aprofundir en el coneixement del mecanisme enzimàtic es van dur a terme estudis de reactivitat mitjançant anàlisi de Hammett: utilitzant una sèrie d'aril-β-D-glucòsids van obtenir-se gràfiques bifàsiques entre log $k_{cat}$ i p$K_a$ de l'aglicó fenòlic. Les gràfiques mostren dependència lineal amb un pendent ($\beta_{lg}$, constant de reacció) igual a -0,8 pels substrats menys reactius (p$K_a$ > 8) i dependència no significativa pels substrats activats (p$K_a$ < 8). Així, d'acord amb el mecanisme en dues etapes de les glicosidases amb retenció de configuració, la formació de l'intermediari glicosil-enzim (glicosidació) és l'etapa limitant pels substrats poc activats (p$K_a$ > 8) mentre que la hidròlisi de l'intermediari (desglicosidació) ho és per aquells molt reactius (p$K_a$ < 8). La caracterització enzimàtica es va completar amb l'anàlisi de la dependència de $k_{cat}$ envers el pH de la reacció: els resultats mostren una doble ionització amb valors de p$K_a$ cinètics d'aproximadament 5,7 i 7,1. D'altra banda, la representació d'Arrhenius, obtinguda amb pNPG com a substrat, va mostrar una relació lineal dins l'interval de temperatura de 30 a 55 ºC; l'energia d'activació calculada és de 49,9 kJ/mol. A temperatures superiors a 60 ºC, l'enzim s'inactiva ràpidament.

En resum, els resultats indiquen que la β-glucosidasa Bgl3 presenta propietats cinètiques i mecanístiques similars a les d'altres enzims de la familia 1 (molt poc estudiats a aquest nivell). Les diferències més rellevants corresponen a les cinètiques de reactivació per azida del mutant E178A i a la diferent especificitat i inhibició per excés de substrat observades en l'enzim natiu. Així, cal concloure que aquests estudis funcionals posen en evidència l'elevada conservació de l'estructura del centre actiu entre els membres de la família i apunten a detalls específics en la seva arquitectura fina com a responsables de les diferències en el comportament cinètic.

# Mechanism of the Family 1 β-Glucosidase from *Streptomyces* sp: Catalytic Residues and Kinetic Studies†

Miquel Vallmitjana,‡ Mario Ferrer-Navarro,‡ Raquel Planell,‡ Mireia Abel,§ Cristina Ausín,§ Enrique Querol,‡ Antoni Planas,*,§ and Josep-Anton Pérez-Pons*,‡

*Institut de Biologia Fonamental "Vicent Villar i Palasí" and Dept Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain, and Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, 08017 Barcelona, Spain*

*Received December 27, 2000; Revised Manuscript Received February 28, 2001*

ABSTRACT: The *Streptomyces* sp. β-glucosidase (Bgl3) is a retaining glycosidase that belongs to family 1 glycosyl hydrolases. Steady-state kinetics with *p*-nitrophenyl β-D-glycosides revealed that the highest $k_{cat}/K_M$ values are obtained with glucoside (with strong substrate inhibition) and fucoside (with no substrate inhibition) substrates and that Bgl3 has 10-fold glucosidase over galactosidase activity. Reactivity studies by means of a Hammett analysis using a series of substituted aryl β-glucosides gave a biphasic plot log $k_{cat}$ vs p$K_a$ of the phenol aglycon: a linear region with a slope of $\beta_{lg} = -0.8$ for the less reactive substrates (p$K_a > 8$) and no significant dependence for activated substrates (p$K_a < 8$). Thus, according to the two-step mechanism of retaining glycosidases, formation of the glycosyl-enzyme intermediate is rate limiting for the former substrates, while hydrolysis of the intermediate is for the latter. To identify key catalytic residues and on the basis of sequence similarity to other family 1 β-glucosidases, glutamic acids 178 and 383 were changed to glutamine and alanine by site-directed mutagenesis. Mutation of Glu178 to Gln and Ala yielded enzymes with 250- and 3500-fold reduction in their catalytic efficiencies, whereas larger reduction ($10^5-10^6$-fold) were obtained for mutants at Glu383. The functional role of both residues was probed by a chemical rescue methodology based on activation of the inactive Ala mutants by azide as exogenous nucleophile. The E178A mutant yielded the β-glucosyl azide adduct (by $^1$H NMR) with a 200-fold increase on $k_{cat}$ for the 2,4-dinitrophenyl glucoside but constant $k_{cat}/K_M$ on azide concentration. On the other hand, the E383A mutant with the same substrate gave the α-glucosyl azide product and a 100-fold increase in $k_{cat}$ at 1 M azide. In conclusion, Glu178 is the general acid/base catalyst and Glu383 the catalytic nucleophile. The results presented here indicate that Bgl3 β-glucosidase displays kinetic and mechanistic properties similar to other family 1 enzymes analyzed so far. Subtle differences in behavior would lie in the fine and specific architecture of their respective active sites.

The Bgl3 β-glucosidase (β-glucoside glucohydrolase, EC 3.2.1.21) from *Streptomyces* sp. QM-B814 is a retaining glycosidase showing an exo-like action pattern by releasing glucose units from the nonreducing end of different β-linked disaccharides and oligosaccharides (*1*). Like other retaining glycosidases, the Bgl3 enzyme is able to perform very efficiently transglycosylation reactions (i.e., transfer of glycosidic residues), catalyzing the formation of β-1,4 or β-1,3 bonds depending on the acceptor (*2*). By means of $^1$H NMR TR-NOE studies it was also determined that *Streptomyces* glucosidase recognizes and bounds only one of the three conformers of thiocellobiose, an analogue of cellobiose which behaves as a competitive inhibitor, in free solution (*3*). Following these structural aims, the protein has been crystallized (*4*), and the 3D structure resolution of its native form and complexes with nonhydrolyzable substrate analogues is currently undertaken (Guasch et al., manuscript in preparation).

In recent years a vast amount of mechanistic information on the enzymatic action of retaining glycosidases has been produced through mutagenesis, enzyme kinetics, inhibition, and X-ray crystallography studies (for reviews see refs *5*−*7*). On the basis of amino acid sequence similarities (*8*), glycosidases have been classified in more that 80 families (CAZY: *ca*rbohydrate-active enzymes at http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html). Enzymes belonging to the same family seem to share common mechanistic features, but given the large number of sequences and protein structures, subtle differences in the fine-tuning of the enzyme activity are expected, thus demanding detailed structural, kinetic, and mechanistic studies of a larger number of glycosidases to evaluate their particular molecular mechanisms.

From a mechanistic point of view, the transfer of glucosyl groups between oxygen nucleophiles catalyzed by β-glucosidases acting with retention of anomeric configuration is

† This work was supported by the Ministerio de Educación y Ciencia (Grants BIO97-0511-C02 and BIO2000-0647-C02 to E.Q. and A.P.) and the Generalitat de Catalunya [Centre de Referència en Biotecnologia (CERBA) to E.Q. and Grup Consolidad GQBB, ref 1999SGR 00335, to A.P.]. M.V. is the recipient of a predoctoral scholarship from CERBA.

* Corresponding authors. A.P.: tel, +34-932038900; fax, +34-932056266; e-mail, aplan@iqs.es. J.-A.P.-P.: tel, +34-935811331; fax, +34-935812011; e-mail, japerez@servet.uab.es.

‡ Universitat Autònoma de Barcelona.

§ Universitat Ramon Llull.

10.1021/bi002947j CCC: $20.00 
Published on Web 04/26/2001

done by means of a double displacement mechanism through the formation of a covalent glycosyl-enzyme intermediate (*5*, *9*). For such a mechanism, two carboxylate residues are generally involved at the enzyme active site: one of them would act as the nucleophile, attacking the anomeric center of the substrate and displacing the leaving group to form the covalent intermediate, and the other as the general acid/base catalyst, providing protonic/deprotonating assistance for the release of the leaving group and the further attack of a water molecule on the covalent intermediate. For family 1 glycosyl hydrolases (*8*), nucleophile and general acid/base residues have been identified by site-directed mutagenesis in the $\beta$-glucosidases Abg from *Agrobacterium faecalis* and SS$\beta$-gly from *Sulfolobus solfataricu*s as Glu358 and Glu170, respectively, for the former enzyme (*10*, *11*) and Glu387 and Glu206, respectively, for the latter enzyme (*40*). Moreover, Glu358 in Abg enzyme was identified as the catalytic nucleophile residue by trapping the covalent $\alpha$-D-glucopyranosyl-enzyme intermediate with the mechanism-based inactivator 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-$\beta$-D-glucopyranoside (2-DFG) and further determination of the labeled residue (*12*). The catalytic mechanism of Agb $\beta$-glucosidase has been extensively studied (*9*, *11*, *13*−*16*), including substrate specificity, pH dependences, kinetic isotope effects, linear free energy relationships, affinity labeling, and interaction with 2-deoxy substrates. This large body of evidence supports the oxocarbocation character of the transition state in both *glycosylation* and *deglycosylation* steps and the covalent nature of the glycosyl-enzyme intermediate, as first proposed by Koshland in the early 50's (*17*). Recently, another family 1 $\beta$-glycosidase from the hyperthermophilic *Pyrococcus furiosus* has been characterized (*18*), showing mechanistic similarities with the mesophilic homologues despite the large difference in their temperature optima. This suggested that structural modifications through evolution have maintained the integrity of the active site structure and specificity of transition state interactions while adapting the overall protein structure to function at notably different temperatures.

According to sequence similarity analysis, Bgl3 $\beta$-glucosidase has also been classified into the family 1 of glycosyl hydrolases, showing identity indexes from 35% to 55% with respect to different members of that family (*1*, *19*). The enzyme is a monomeric protein with molecular mass of 52.6 kDa and an isoelectric point of 4.4; it has a pH optimum for activity of 6.5, and the optimum temperatute for activity is 50 °C. Following sequence alignment and molecular modeling, we proposed the residues Glu383 and Glu178 as candidates to play the role of nucleophile and acid/base catalyst, respectively. In this paper we report the kinetic properties of the wild-type Bgl3 and mutants at the proposed catalytic residues with the aim of comparing the mechanistic details with other family 1 $\beta$-glucosidases and to provide further evidence on the conservation of active site structures among members of the same enzyme family according to the current classification of glycosyl hydrolases.

## MATERIALS AND METHODS

*Reagents and Substrates.* Phenyl, 2-nitrophenyl, and 4-nitropheyl $\beta$-D-glucopyranoside and *p*-nitrophenyl $\beta$-D-glycosides of xylose, fucose, and galactose as well as all buffer chemicals were obtained from Sigma Chemical Co. 4-Bromo-, 3-nitro-, 3,5-dinitro-, 3,4-dinitro-, and 2,3-dinitrophenyl $\beta$-D-glucopyranosides were prepared by phase transfer-catalyzed glycosylation of $\alpha$-bromoglucose peracetate with the corresponding phenol as reported in Dess et al. (*20*), followed by de-O-acetylation in MeONa/MeOH. 4-Methylumbelliferyl $\beta$-D-glucopyranoside was synthesized as reported in Malet et al. (*21*) and 2,4-dinitrophenyl $\beta$-D-glucopyranoside by the one-step procedure of Sharma et al. (*22*).

*Cloning, Expression, and Purification of Bgl3 Enzymes from Escherichia coli.* The wild-type (wt) $\beta$-glucosidase gene from *Streptomyces* sp. QM-B814, including its downstream region which may function as transcription terminator since it contains two inverted repeats (CGGTGCGGCAC and GGCCCGCCCCCG starting at 8 and 37 nucleotides from the stop codon, respectively), was cloned into the T7 polymerase expression vector pET-21d(+) (Novagen) and overexpressed in *E. coli* strain BL21(DE3) (Novagen). To simplify the protein purification protocol, a His-tag coding sequence was fused to the start codon (GTG) of the Bgl3 reading frame, obtaining the recombinant plasmid pET21-HBG3 which yields a protein with an extended N-terminus formed by the sequence MHHHHHHGIH and a deduced molecular mass of 53.6 kDa.

Each expression experiment was started by transforming competent *E. coli* BL21(DE3) cells with pET21-derived plasmids containing wt or mutant Bgl3 genes. One transformant colony grown in LB solid medium, containing ampicillin (100 $\mu$g/mL), was picked, inoculated into 3 mL of 2× YT broth + ampicillin (200 $\mu$g/mL) (*23*), and shaken overnight at 37 °C to an $OD_{595}$ of about 0.6; aliquots of this culture were used to inoculate larger volumes of the same growth medium (usually 200 mL) that were shaken (300 rpm) for 8−10 h at 37 °C. Following such a procedure, protein expression was achieved constitutively without addition of isopropyl $\beta$-thiogalactopyranoside (IPTG). The cells were harvested and disrupted essentially according to the protocol described in the pET System Manual (Novagen); the sonication step was omitted. For protein purification, cell-free extracts were centrifuged (39000*g*, 30 min at 4 °C), and the supernatant (soluble fraction) was applied onto a 5 mL HiTrap Chelating Sepharose (Pharmacia) column previously charged with $Ni^{2+}$ as a metal ion. Chromatography was carried out according to the manufacturer's specifications, and Bgl3-containing fractions were pooled, concentrated, and desalted by ultrafiltration at 4 °C using 50 mM sodium phosphate (pH 7.0) as desalting buffer. Using this procedure homogeneous (>99%) protein samples were obtained, as judged by SDS−polyacrylamide gel electrophoresis (SDS−PAGE) and Coomassie Brilliant Blue R-250 staining. Protein concentration was determined by the dye-binding method of Bradford (*24*) using BSA as a standard. Pure enzymes were kept at 4 °C for short periods or stored at −20 °C in the presence of 45% glycerol for prolonged periods. A protein extinction coefficient of 107 000 $M^{-1}$ $cm^{-1}$ at 280 nm was calculated for the wt Bgl3 according to ref *25*.

*Site-Directed Mutagenesis.* Different fragments cloned in pUC118, previously obtained for the *bgl3* gene sequencing (*1*), were used as templates for mutagenic polymerase chain reactions (SDM-PCR) according to the method reported by Juncosa et al. (*26*). The first PCR used the mutagenic primers and the Reverse Universal Primer flanking the 5′-end of the

*bgl3* gene fragment including the target triplet. Mutagenic primers were as follows (mismatches are underlined): E383A, 5′-CTGGTCATCACCGCGAACGGCGCCGC-3′; E383Q, 5′-CGCTGGTCATCACCCAGAACGGCGCCGCC-3′;E178Q, 5′-GGACCACCCTCAACCAGCCCTGGTGCAGC-3′;E178A, 5′-GGACCACCCTCAACGCGCCCTGGTGCAGC-3′. The first PCR product was purified by agarose gel electrophoresis and used as a primer in the second PCR along with the Forward Universal Primer. Mutated fragments were purified by agarose gel electrophoresis, cloned in pUC118 vector, and completely resequenced before their substitution for the corresponding fragment in the wild-type pET21-HBG3.

*Equilibrium Urea Denaturation.* Unfolding of wt and mutated Bgl3 was monitored by fluorescence spectroscopy in a Perkin-Elmer LS 50 spectrofluorometer, with excitation at 280 nm (2.5 nm slit) and the emission spectra being recorded from 300 to 550 nm (4 nm slit) and measured at 340 nm, in a thermostated cuvette at 25 °C. Protein samples in 50 mM sodium phosphate buffer, pH 7.0, were diluted to a concentration of 40 μg/mL in degassed urea solutions in the same buffer and incubated 8−10 h at room temperature before fluorescence data were collected. Equilibrium denaturation data were analyzed using the Clarke and Fersht method (*27*) as previously described by Pons et al. (*28*) by fitting the fluorescence vs urea concentration data to

$$F = (a_F + b_F[\mathrm{D}] + (a_U + b_U[D]) \exp(m([\mathrm{D}] - [\mathrm{D}]_{50\%})/RT))/(1 + \exp(m([\mathrm{D}] - [\mathrm{D}]_{50\%})/RT)) \quad (1)$$

where $F$ is the measured fluorescence, $a_F$ and $a_U$ are the intercepts and $b_F$ and $b_U$ the slopes of the baselines at low (F) and high (U) denaturant concentration, [D] is the denaturant (urea) concentration, $[D]_{50\%}$ is the concentration of denaturant at which 50% of the protein is unfolded, and $m$ ($=\delta\Delta G_U/\delta[\mathrm{D}]$) is the slope of the transition.

*Kinetic Characterization. (A) General.* β-Glucosidase activity was determined by a continuous assay monitoring the release of *p*-nitrophenol from *p*-nitrophenyl β-D-glucopyranoside (pNPG) at 400 nm in a thermostated spectrophotometer. Enzyme reactions were performed in 50 mM sodium phosphate buffer, pH 6.5 or 7.0, at 50 °C using different pNPG concentrations (0.01−20 mM). The molar absorption coefficients at 400 nm ($\Delta\epsilon = \epsilon_{p-\mathrm{nitrophenol}} - \epsilon_{\mathrm{substrate}}$) were 5486 $M^{-1}\cdot cm^{-1}$ at pH 6.5 and 9510 $M^{-1}\cdot cm^{-1}$ at pH 7.0 at 50 °C. The protein concentrations in the reaction mixtures were 2−20 nM for the wild-type enzyme and 0.05−2 μM for the mutants. Rates were obtained from the initial slopes (2−5 min reaction time for the wild-type enzyme and up to 45 min for the mutants), and kinetic constants were calculated from data fitted to the Michaelis−Menten equation or to a substrate inhibition model ($v = k_{cat}[\mathrm{E}]_0[\mathrm{S}]/(K_m + [\mathrm{S}] + [\mathrm{S}]^2/K_i)$) by means of nonlinear regression analysis. Activity against other *p*-nitrophenyl sugars was analyzed following the assay described above.

*(B) pH Profile.* The pH dependence of enzyme activity was studied using pNPG (0.5 mM) as a substrate in 50 mM phosphate buffer, $I = 0.15$ M with added KCl, at 40 °C. The stability of the enzyme at each pH was first determined, and activity was calculated at pH values at which the enzyme was stable for at least 5 min. The pH-dependent molar extinction coefficients of *p*-nitrophenol at 378 nm (for reaction at pH ≤5.5) and 400 nm (for reactions at pH >5.5) were accurately measured at the conditions used for the kinetic assays. Values of $\epsilon_{378}$ and $\epsilon_{400}$ at different pH values were fitted by nonlinear regression to the equation

$$\epsilon_{\mathrm{pNP-OH},\lambda} = (\epsilon_{\mathrm{NPH},\lambda} + \epsilon_{\mathrm{NP}^-,\lambda} \times 10^{\mathrm{pH}-\mathrm{p}K_a})/(1 + 10^{\mathrm{pH}-\mathrm{p}K_a}) \quad (2)$$

where NPH and $NP^-$ are the un-ionized and ionized forms of *p*-nitrophenol, respectively, and $\epsilon_{\mathrm{pNP-OH},\lambda}$ is the measured molar extintion coeficient at each $\lambda$. The adjusted parameters were $\epsilon_{\mathrm{NPH},378} = 992\ M^{-1}\ cm^{-1}$, $\epsilon_{\mathrm{NP}^-,378} = 13\,695\ M^{-1}\ cm^{-1}$, $\epsilon_{\mathrm{NPH},400} = 114\ M^{-1}\ cm^{-1}$, $\epsilon_{\mathrm{NP}^-,378} = 18\,131\ M^{-1}cm^{-1}$, and $pK_a = 6.92$.

*(C) Hammett Analysis.* Kinetic constants for the enzyme-catalyzed hydrolysis of a series of aryl β-D-glucosides were determined by measuring the release of the aglycon by UV spectrophotometry in 50 mM phosphate buffer, pH 6.5, $I = 0.15$ M with added KCl, at 30 °C. Molar extinction coefficients of the aryl β-glucosides and the aglycons (free phenols) were determined at the appropriate wavelength under the same conditions [wavelength and $\Delta\epsilon$ ($=\epsilon_{\mathrm{phenol}} - \epsilon_{\mathrm{substrate}}$) are listed in Table 3].

*(D) Chemical Rescue.* Kinetics of wt and mutant enzymes in the presence of sodium azide (0−100 mM) were evaluated for *p*-nitrophenyl β-D-glucopyranoside (pNPG) and 2,4-dinitrophenyl β-D-glucopyranoside (2,4-DNPG)[1] substrates by monitoring the release of the phenol aglycon in 50 mM phosphate buffer, pH 6.5, at 50 °C for pNPG ($\lambda = 400$ nm, $\Delta\epsilon = 5486\ M^{-1}\ s^{-1}$), and 30 °C for 2,4-DNPG ($\lambda = 400$ nm, $\Delta\epsilon = 2178\ M^{-1}\ s^{-1}$). Substrate and enzyme concentrations are given in Table 4.

*(E) NMR Monitoring.* $^1$H NMR monitoring of enzymatic reactions was performed following the procedure reported in Viladot et al. (*29*) on a Varian Gemini 300 spectrometer at 30 °C in 9 mM sodium phosphate buffer at pD 7.0 (adjusted with 1 N NaOD/$D_2O$). Spectra were recorded at intervals of 15 min during a total time of 10−24 h. Enzyme and nucleophile concentrations were (a) E178A [enzyme] = 300 nM, [azide] = 100 mM, [substrate] = 3.75 mM and (b) E383A [enzyme] = 800 nM, [azide] = 200 mM, [substrate] = 3.2 mM.

## RESULTS AND DISCUSSION

*Mutagenesis and Expression.* To undertake both mutational and structural studies with the Bgl3 glucosidase, the availability of a quick protein expression and purification procedure was the first pitfall to be overcome. The Bgl3 encoding gene was initially cloned in *Streptomyces lividans* where it expresses intracellularly from the own promoter (*1*); such a promoter is not functional in *E. coli*. Moreover, pure enzyme preparations from the streptomycete host involve from three to seven purification steps, depending on the volume of culture processed, with the concomitant decrease in protein yields. Thus, a series of attempts to express Bgl3

[1] Abbreviations: pNPGal, *p*-nitrophenyl β-D-galactoside; pNPGlu, *p*-nitrophenyl β-D-glucoside; pNPFuc, *p*-nitrophenyl β-D-fucoside; pNPXyl, *p*-nitrophenyl β-D-xyloside; PG, phenyl β-D-glucoside; 4-BrPG, 4-bromophenyl β-D-glucoside; 3-NPG, 3-nitrophenyl β-D-glucoside; MUG, 4-methylumbelliferyl β-D-glucoside; 2-NPG, 2-nitrophenyl β-D-glucoside; 4-NPG or pNPG, 4-nitrophenyl β-D-glucoside; 3,5-DNPG, 3,5-dinitrophenyl β-D-glucoside; 3,4-DNPG, 3,4-dinitrophenyl β-D-glucoside; 2,3-DNPG, 2,3-dinitrophenyl β-D-glucoside; 2,4-DNPG, 2,4-dinitrophenyl β-D-glucoside.

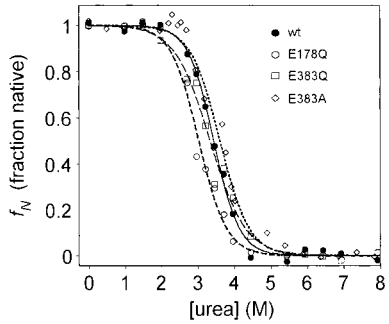



FIGURE 1: Equilibrium urea denaturation curves for the wild-type and mutant $\beta$-glucosidases. Fluorescence intensity values were fitted to eq 1, and the adjusted parameters are given in Table 1. Normalized curves are plotted as the fraction native $f_N = (F - (a_U + b_U[D]))/(a_F - a_U + [D](b_F - b_U))$, where $F$ is the measured fluorescence intensity and $a_U$, $a_F$, $b_U$, and $b_F$ are the adjusted parameters from eq 1.

in *E. coli* were carried out, including extracellular expression by fusion to several signal peptides from *E. coli* and *Bacillus licheniformis*. Among other vectors successfully tested [i.e., pUCBM21 (Boehringer Mannheim) expressing Bgl3 from *lac* promoter], the pET system (Novagen) and N-terminal fusion of a His tag have allowed to obtain very high yields of soluble enzyme (150−200 mg/L of *E. coli* culture) following a one-step purification protocol. In contrast, the His-tag fusion to the C-terminal end of Bgl3 glucosidase was absolutely deleterious for its expression in a soluble form as was any attempt of secretion by fusing different signal peptides to the N-terminal end. The recombinant Bgl3 carrying the N-terminal extension (see Materials and Methods) obtained from *E. coli* did not differ significantly from the enzyme purified from *S. lividans* comparing activity−pH profiles, substrate specificity, kinetic constants, and thermostability (not shown).

The choice of the amino acid residues essential for catalysis was based on sequence alignment between Bgl3 and different members of family 1 glycosyl hydrolases, taking into account that glutamic acid residues 358 and 170 had been previosly identified as the nucleophile and general acid/base catalyst, respectively, in the Abg enzyme (*10*−*12*). The corresponding residues in Bgl3 are Glu383 and Glu178, which are found in the highly conserved motifs I-T-E-N-G-A and N-E-P-W, respectively. On the basis of the first 3D structure of a glucosidase belonging to the family 1 glycosyl hydrolases (*30*), a model structure of Bgl3 was obtained and used to verify the positioning of the residues chosen upon alignment. On the modeled structure, the putative role of Glu383 and Glu178 residues was reinforced, being located in the inner region of the active site pocket and presenting a distance of about 0.4 nm between their respective carboxylate groups.

Glu383 and Glu178 were substituted by the isosteric glutamine (charge removal) and by alanine (side chain and charge removal). The mutant Bgl3 enzymes were obtained by site-directed mutagenesis and expressed and purified following the procedure described for the wild type.

*Unfolding.* To ascribe the changes in kinetic behavior to the introduced mutations, the proper folding or stability of the mutants was evaluated by equilibrium urea denaturation. Thus, unfolding was monitored by measuring the dependence of fluorescence intensity on urea concentration (Figure 1) and fitting the fluorescence vs urea concentration to eq 1 (Table 1). Considering that $[D]_{50\%}$ is the concentration of denaturant at which 50% of the protein is unfolded, the mean $[D]_{50\%}$ value between mutant enzymes, as calculated from the fluorescence curves, was 3.3 M urea (with a standard deviation of ±0.3). Such a value compares well with $[D]_{50\%}$ calculated for wild-type Bgl3, i.e., 3.44 M urea, as do the denaturation curves from 0 to 8 M urea (Figure 1). These results support a folded structure for the mutant proteins, and hence, changes of the catalytic properties can be assigned to the replacement of the target amino acid residues.

Table 1: Equilibrium Urea Denaturation Data for wt and Mutant $\beta$-Glucosidases[a]

| mutant | m (kcal·mol$^{-1}$·M$^{-1}$) | $[D]_{50\%}$ (M) | $\Delta G^{H_2O}$ [b] (kcal·mol$^{-1}$) |
|---|---|---|---|
| wild type | 1.69 | 3.44 ± 0.02 | 5.8 |
| E178Q | 1.56 | 3.03 ± 0.04 | 4.7 |
| E383Q | 1.23 | 3.34 ± 0.04 | 4.1 |
| E383A | 1.48 | 3.61 ± 0.04 | 5.3 |

[a] Conditions: 50 mM phosphate, pH 7.0, 0−8 M urea, 40 $\mu$g/mL enzyme. [b] $\Delta G^{H_2O} = m[D]_{50\%}$.

Table 2: Michaelis−Menten Parameters of *Streptomyces* sp. $\beta$-Glucosidases (wt and Mutants) with Cellobiose and *p*-Nitrophenyl $\beta$-Glycoside Substrates[a]

| mutant | substrate | $K_M$ (mM) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_M$ (M$^{-1}$ s$^{-1}$) |
|---|---|---|---|---|
| wt | cellobiose | 4.1 | 35.6 | $8.60 \times 10^3$ |
| | PNPGlu | 0.15 | 28.4 | $1.90 \times 10^5$ |
| | PNPFuc | 0.14 | 37.1 | $2.65 \times 10^5$ |
| | PNPGal | 7.3 | 118 | $1.61 \times 10^4$ |
| | PNPXil | 3.0 | 0.63 | $2.11 \times 10^2$ |
| E178Q | PNPGlu | 0.02 | 0.015 | $7.77 \times 10^2$ |
| | PNPFuc | 0.11 | 1.08 | $9.90 \times 10^3$ |
| | PNPGal | 4.5 | 0.66 | $1.50 \times 10^2$ |
| E178A | PNPGlu | 0.65 | 0.036 | $5.52 \times 10^1$ |
| | PNPFuc | 0.22 | 0.011 | $4.86 \times 10^1$ |
| | PNPGal | 0.12 | 0.004 | $3.65 \times 10^1$ |
| E383Q | PNPGlu | 0.20 | $\approx 3 \times 10^{-5}$ | $1.6 \times 10^{-1}$ |
| | PNPFuc | 0.26 | $\approx 2 \times 10^{-5}$ | $6.0 \times 10^{-2}$ |
| | PNPGal | 0.04 | $\approx 1 \times 10^{-4}$ | 4.0 |
| E383A | PNPGlu | 0.03 | $\approx 4 \times 10^{-5}$ | 1.4 |
| | PNPFuc | 0.23 | $\approx 2 \times 10^{-4}$ | $9.2 \times 10^{-1}$ |
| | PNPGal | 0.04 | $\approx 8 \times 10^{-4}$ | $2.0 \times 10^1$ |

[a] Conditions: 50 mM phosphate buffer, pH 6.5, 50 °C (for wt, E178Q, and E178A), pH 7.0, 30 °C (for E383Q and E383A), [enzyme] = 0.02 $\mu$M (wt), 0.075−0.2 $\mu$M (E178Q), 0.048−0.28 $\mu$M (E178A), 1.18 $\mu$M (E383Q), and 0.54 $\mu$M (E383A). Protein concentrations were determined by the Bradford assay using BSA as a standard.

*Substrate Specificity.* Kinetics for the wild-type and mutant $\beta$-glucosidases were evaluated with different *p*-nitrophenyl $\beta$-glycosides as summarized in Table 2. The wt enzyme has a broad substrate specificity which is common among glycosyl hydrolases from family 1, such as $\beta$-glucosidases from *Agrobacterium faecalis* (*13*), *Bacillus polymyxa* (*31*), *Thermotoga maritima* (*32*), *Sulfolobus solfataricus* (*33*), and *Pyrococcus furiosus* (*18*). Glucoside and fucoside substrates show the highest catalytic efficiency (in terms of $k_{cat}/K_M$), but a significant substrate inhibition is obtained for the *p*-nitrophenyl $\beta$-glucoside as opposed to the fucoside (Figure 2). The $k_{cat}$ value for the galactoside (pNPGal) substrate is 4-fold higher than that for the glucoside (pNPGlu), but the lower $K_M$ for the latter results in an overall 10-fold glucosidase over galactosidase activity in terms of $k_{cat}/K_M$.

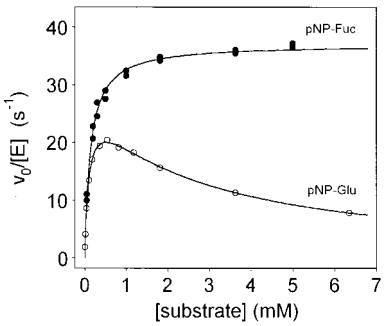



FIGURE 2: Steady-state kinetics for the enzyme-catalyzed hydrolysis of *p*-nitrophenyl $\beta$-D-glucoside (pNPGlu) and *p*-nitrophenyl $\beta$-L-fucoside (pNPFuc) at 50 °C in 50 mM phosphate buffer, pH 6.5.

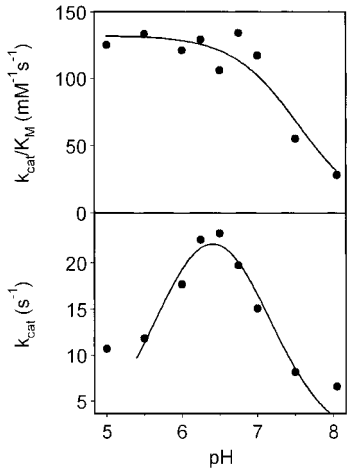



FIGURE 3: pH dependence of the kinetic parameters of wild-type *Streptomyces* sp. $\beta$-glucosidase with *p*-nitrophenyl $\beta$-D-glucoside: (a, bottom) pH profile of $k_{cat}$; (b, top) pH profile of $k_{cat}/K_M$.

The lowest efficiency among the substrates assayed corresponds to the xyloside (pNPXyl) with a 1000-fold decrease in $k_{cat}/K_M$ as compared to the pNPGlu substrate, indicating that the interactions between the 5-hydroxymethyl group and the binding pocket of the enzyme have a stabilizing effect on the transition state. The magnitude of this interaction can be estimated according to (*18*, *34*)

$$\Delta\Delta G^{\ddagger} = -RT \ln[(k_{cat}/K_M)_{xyl}/(k_{cat}/K_M)_{glu}] \qquad (3)$$

where $\Delta\Delta G^{\ddagger}$ is the change in the activation free energy between both substrates and $(k_{cat}/K_M)_{xyl}$ and $(k_{cat}/K_M)_{glu}$ are the catalytic efficiencies for the pNPXyl and pNPGlu substrates, respectively. A value of 18.5 kJ·mol$^{-1}$ is lost upon removal of the 5-hydroxymethyl group, comparable to the 16.3 kJ·mol$^{-1}$ for the *A. faecalis* (*13*) and the 10 kJ·mol$^{-1}$ for the *P. furiosus* (*18*) $\beta$-glucosidases.

*pH and Temperature Dependence.* The steady-state values of $k_{cat}$, $K_M$, and $k_{cat}/K_M$ for the $\beta$-glucosidase-catalyzed reaction of pNPGlu were measured over the pH range of 5−8 at 40 °C and constant ionic strength (0.15 M) (Figure 3). The pH dependence of $k_{cat}$ follows a double ionization curve, with kinetic p$K_a$ values of aproximately 5.7 and 7.1. $k_{cat}$ parameters at low and high pH could not be accurately measured due to large increases in $K_M$ values. Moreover, enzyme precipitation was apparent at pH <5. In the pH range studied, a single ionization curve was obtained for $k_{cat}/K_M$, with a kinetic p$K_a$ in the basic limb of 7.5. Since the pH profile on $k_{cat}/K_M$ reflects ionization constants in the free enzyme, and on $k_{cat}$ in the E·S complex, the low p$K_a$ in the free enzyme (not observed in the $k_{cat}/K_M$ plot) should be below 5, which is shifted to 5.7 in the E·S complex. On the other hand, catalysis depends on a protonated group with a p$K_a$ of 7.1−7.5 (on the E·S complex and free enzyme, respectively). Comparison with other family 1 $\beta$-glucosidases shows a similar behavior, mainly for the protonated group with high p$K_a$: 7.0−7.4 for the *P. furiosus* $\beta$-glucosidase (*18*) and 7.6−8.1 for the *A. faecalis* enzyme (*13*), indicating that the specific hydrogen bonds that occur between the catalytic amino acids and other active site residues are similar in these enzymes.

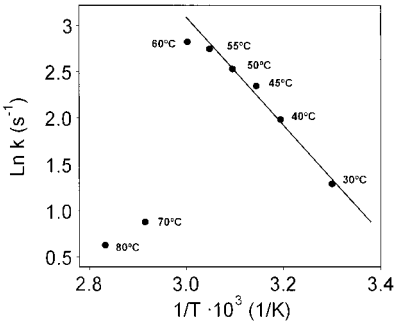
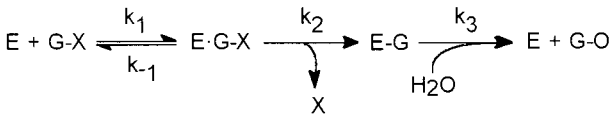



FIGURE 4: Temperature dependence of the $\beta$-glucosidase activity on *p*-nitrophenyl $\beta$-D-glucoside substrate in 50 mM phosphate buffer, pH 6.5, [pNPGlu] = 588 $\mu$M, [enzyme] = 5.6 nM, and temperature range 30−80 °C.

The *Streptomyces* $\beta$-glucosidase shows a linear Arrhenius plot for the reaction with the pNPGlu substrate in the temperature range from 30 to 55 °C (Figure 4). The calculated activation energy ($E_a$) of 49.9 kJ·mol$^{-1}$ is of the same magnitude than those determined for other family 1 $\beta$-glucosidases (*18*). At temperatures higher than 60 °C, the enzyme is rapidly inactivated.

*Enzyme Mechanism: Structure−Reactivity Studies.* As a retaining glycosidase, the mechanism of Bgl3 involves a two-step process (*5*, *6*). In the first step (*glycosylation*) the amino acid residue acting as a general acid protonates the glycosidic oxygen with concomitant C−O breaking of the scissile glycosidic bond, while the deprotonated carboxylate functioning as a nucleophile attacks the anomeric center to give a covalent glycosyl-enzyme intermediate (E−G). The second *deglycosylation* step involves the attack of a water molecule assisted by the conjugate base of the general acid to render the free sugar with overall retention of configuration (G−OH). Both steps proceed via transition states with considerable oxocarbenium ion character.




In this mechanism, $k_{cat} = k_2k_3/(k_2 + k_3)$, $K_M = k_3(k_2 + k_{-1})/k_1(k_3 + k_2)$, and $k_{cat}/K_M = k_1k_2/(k_2 + k_{-1})$.

A series of substituted aryl $\beta$-glucosides were used as substrates of the $\beta$-glucosidase-catalyzed reaction to evaluate the effect of the aglycon leaving group on reaction rates. Kinetic data are presented in Table 3. All substrates showed substrate inhibition at high substrate concentration, and $k_{cat}$ and $K_M$ values were obtained by fitting the initial velocities vs substrate concentration to the Michaelis−Menten equation

Table 3: Kinetic Parameters of Aryl $\beta$-D-Glucosides with *Streptomyces* sp. $\beta$-Glucosidase at 30 °C[a]

| substrate | concn range (mM) | $\Delta\epsilon$ ($M^{-1}\cdot cm^{-1}$) [$\lambda$ (nm)][b] | p$K_a$ (phenol) | $k_{cat}$ ($s^{-1}$) | $K_M$ (mM) | $K_I$ (mM) | $k_{cat}/K_M$ ($M^{-1}\cdot s^{-1}$) |
|---|---|---|---|---|---|---|---|
| PG | 0.04−9.3 | 1414 [270] | 9.99 | 0.66 ± 0.08 | 2.3 ± 0.4 | 16 ± 7 | $2.83 \times 10^2$ |
| 4-BrPG | 0.06−4.8 | 909 [288] | 9.34 | 3.32 ± 0.12 | 0.47 ± 0.03 | 28 ± 10 | $7.11 \times 10^3$ |
| 3-NPG | 0.01−1.6 | 224 [385] | 8.39 | 15.8 ± 2.10 | 0.24 ± 0.05 | 1.2 ± 0.3 | $6.52 \times 10^4$ |
| MUG | 0.001−1.3 | 1986 [355] | 7.53 | 11.7 ± 0.30 | 0.058 ± 0.004 | 4.6 ± 0.8 | $2.04 \times 10^5$ |
| 2-NPG | 0.002−1.0 | 1459 [379] | 7.22 | 7.72 ± 0.90 | 0.023 ± 0.017 | 0.18 ± 0.05 | $3.43 \times 10^5$ |
| 4-NPG | 0.002−1.7 | 5088 [400] | 7.18 | 8.14 ± 0.40 | 0.10 ± 0.01 | 1.4 ± 0.1 | $7.85 \times 10^4$ |
| 3,5-DNPG | 0.01−1.8 | 87 [400] | 6.69 | 20.5 ± 0.70 | 0.054 ± 0.005 | 11 ± 4 | $3.78 \times 10^5$ |
| 3,4-DNPG | 0.003−1.2 | 8953 [400] | 5.36 | 18.2 ± 1.0 | 0.064 ± 0.007 | 6.4 ± 0.6 | $2.87 \times 10^5$ |
| 2,3-DNPG | 0.003−0.4 | 2475 [420] | 4.96 | 11.7 ± 0.7 | 0.015 ± 0.002 | 2.1 ± 1.5 | $7.70 \times 10^5$ |
| 2,4-DNPG | 0.001−0.1 | 2178 [400] | 3.96 | 30.4 ± 1.0 | 0.008 ± 0.001 | 0.23 ± 0.04 | $3.83 \times 10^6$ |

[a] Conditions: 50 mM phosphate buffer, pH 6.5, $I = 0.15$ M with added KCl, 30 °C. [b] $\lambda$, wavelength at which the release of the aglycon was monitored spectrophotometrically; $\Delta\epsilon$ values ($=\epsilon_{phenol} - \epsilon_{substrate}$) were determined at the given wavelength in the same buffer and temperature used in the kinetic runs. Protein concentrations were determined by the Bradford assay using BSA as a standard.

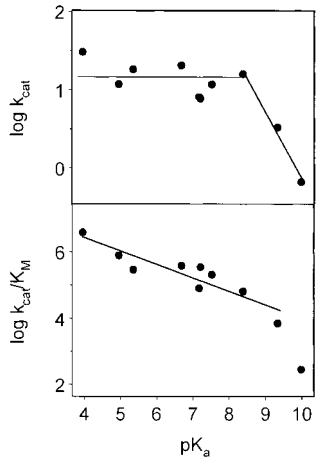



FIGURE 5: Hammett relationship for the $\beta$-glucosidase-catalyzed hydrolysis of aryl $\beta$-glucosides. Data are presented in the form of a Brønsted plot (log $k_{cat}$ versus p$K_a$ of the aglycon phenol). Kinetic data and experimental conditions are given in Table 3.

with substrate inhibition model. The effect of substrate reactivities on $k_{cat}$ and $k_{cat}/K_M$ was evaluated by means of a Hammett analysis, the electronic effects of the phenol substituents being expressed as the p$K_a$ of the free phenol released upon enzymatic hydrolysis. Therefore, the Hammett equation is written in the form of the Brønsted equation

$$\log k = \text{constant} + \beta_{lg} pK_a$$

where $k$ is $k_{cat}$ or $k_{cat}/K_M$ and $\beta_{lg}$ is the reaction constant reflecting the sensitivity of the reaction to the electronic effect of the substituents. The Hammett relationships between reaction rates and leaving group ability (as p$K_a$ of the phenols) are presented in Figure 5. A biphasic plot of log $k_{cat}$ vs p$K_a$ was obtained, which corresponds to a change in the rate-determining step from less reactive to more reactive substrates. The plot is linear with a slope (Brønsted coefficient $\beta_{lg}$) of −0.8 for substrates with poor leaving ability (p$K_a$ > 8), indicating that the first *glycosylation* step is rate limiting and $k_{cat}$ is proportional to the reactivity of the substrate. Glycosides with good leaving groups (p$K_a$ < 8) show no significant dependence of log $k_{cat}$ on p$K_a$, the aglycon-independent *deglycosylation* step now being rate limiting. This biphasic behavior is similar to that observed for the *A. faecalis* and *P. furiosus* $\beta$-glucosidases (*13*, *18*), with close reaction constants ($\beta_{1g} = -0.7$) for the less reactive substrates; the formation of the glycosyl-enzyme intermediate is rate limiting, and there is a large degree of bond cleavage with charge development at the transition state for substrates with poor leaving groups, whereas hydrolysis of the glycosyl-enzyme intermediate becomes rate limiting for activated substrates.

The Hammett plot for the second-order rate constant $k_{cat}/K_M$ (Figure 5) is linear and also shows a concave downward curvature at high p$K_a$ values. Since $k_{cat}/K_M$ contains rate constants up to the first irreversible step (which is *glycosylation*), the curvature cannot be related to a change in rate-limiting *glycosylation* to *deglycosylation*, but rather it may reflect a change in the rate of enzyme−substrate association ($k_1$) relative to the rate of *glycosylation* ($k_2$).

*Catalytic Residues.* Mutation of the proposed catalytic residues, Glu178 and Glu383, to Gln and Ala yields less active to inactive enzymes (Table 2). For the Glu178 residue, assigned to the general acid/base on the basis of sequence similarities, mutation to Gln gave a 250-fold reduction of $k_{cat}/K_M$ (pNPGlu substrate), whereas mutation to Ala produced a 3500-fold decrease in catalytic efficiency. Larger effects are obtained for mutants at Glu383, proposed as the catalytic nucleophile, with $10^6-10^5$-fold reductions in $k_{cat}/K_M$ for the E383Q and E383A mutants.

*Chemical Rescue of Inactive Mutants.* The functional role of both residues in the enzyme mechanism was assessed by a chemical rescue methodology with an exogenous nucleophile as applied to a number of glycosidases [i.e., $\beta$-glucosidases (*11*, *35*, *36*), 1,4-$\beta$-exoglucanase (*37*), 1,3-1,4-$\beta$-glucanase (*29*, *38*), $\beta$-galactosidase (*39*)]. For a retaining $\beta$-glycosidase, removal of the general acid/base residue will have little or no effect on the first *glycosylation* step for an activated substrate (2,4-DNP glycoside) that does not require general acid assistance, but the reaction will stop at the *deglycosylation* step. The glycosyl-enzyme intermediate will be slowly hydrolyzed in the absence of the general base to assist the attack of a water molecule. Addition of an exogenous nucleophile (such as azide) that does not require general base assistance may attack the glycosyl-enzyme intermediate instead of a water molecule with the overall effect of enzyme reactivation. Since the exogenous nucleophile will operate on the *deglycosylation* step, the reaction will yield the $\beta$-glycosyl product. On the other hand, mutation of the nucleophile should render an inactive mutant where the first *glycosylation* step has been drastically slowed because of the absence of the nucleophilic residue to form the glycosyl-enzyme intermediate. In this case, a cavity has

Table 4: Kinetic Parameters for Wild-Type and Mutant *Streptomyces* sp. $\beta$-Glucosidases in the Absence and Presence of Sodium Azide (Chemical Rescue of Inactive Mutants)[a]

| substrate | enzyme | azide (mM) | $K_M$ ($\mu$M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_M$ ($M^{-1}s^{-1}$) |
|---|---|---|---|---|---|
| pNPG | wt | | 150 | 28.4 | $1.90 \times 10^5$ |
| | E178A | 0 | 648 | 0.036 | $5.52 \times 10^1$ |
| | | 50 | 1436 | 0.12 | $8.18 \times 10^1$ |
| | E178Q | 0 | 19.7 | 0.015 | $7.77 \times 10^2$ |
| | | 50 | 575 | 1.38 | $2.40 \times 10^3$ |
| | E383A[b] | 0 | 26 | $\approx 4 \times 10^{-5}$ | 1.4 |
| | | 255 | 36 | $\approx 4 \times 10^{-5}$ | 1.1 |
| 2,4-DNPG | wt | | 8.0 | 30.4 | $3.83 \times 10^6$ |
| | E178A | 0 | 0.3 | 0.022 | $7.4 \times 10^4$ |
| | | 50 | 1.8 | 4.14 | $2.3 \times 10^6$ |
| | E178Q | 0 | 5.4 | 0.076 | $1.4 \times 10^4$ |
| | | 50 | 3.2 | 2.04 | $6.4 \times 10^5$ |
| | E383A | 0 | 27 | $3.5 \times 10^{-4}$ | $1.3 \times 10^1$ |
| | | 255 | 168 | $3.1 \times 10^{-3}$ | $1.9 \times 10^1$ |
| | | 1000 | 990 | $4.7 \times 10^{-2}$ | $4.8 \times 10^1$ |

[a] Conditions: 50 mM phosphate buffer, pH 6.5, $I = 0.15$ M with added KCl, 50 °C (for pNPG substrate), 30 °C (for 2,4-DNPG substrate). [b] Parameters determined at 30 °C, pH 7.0. Protein concentrations were determined by the Bradford assay using BSA as a standard.

been created in the active site on the α-face (Glu to Ala mutation) which can accommodate a small exogenous nucleophile such as azide. The enzyme might be reactivated through a different mechanism, following a single inverting displacement to give the α-glycosyl product. Therefore, if reactivation occurs, the stereochemistry of the new glycosyl azide product formed upon chemical rescue of both alanine mutants will indicate which residue acts as the general acid/base and which operates as the nucleophile in the catalytic mechanism of the wild-type enzyme.

Addition of sodium azide as exogenous nucleophile to the reaction of the inactive alanine mutants at each catalytic residue with an activated substrate (2,4-DNPGlu) restored enzyme activity. The enzymatic reactions were monitored by $^1$H NMR spectroscopy (in $D_2O$ at pD 7.0 and 30 °C) for structure determination of the final products. The reactions took 1−3 days to completeness at the low enzyme concentrations used in the assays. The reaction of E178A in the presence of 100 mM azide yielded the $\beta$-glucosyl azide product: a doublet at $\delta$ 4.75 ppm ($J = 8$ Hz), which corresponds to H-1 of the $\beta$-anomer, develops at the same rate at which the doublet corresponding to H-1 of the substrate ($\delta$ 5.45 ppm, $J = 7.5$ Hz) disappears. No hydrolysis products are detected when the final spectrum is compared to that of the wild-type reaction with the same substrate, where the $\beta$ ($\delta$ 4.66 ppm) and α ($\delta$ 5.23 ppm) anomers of glucose (hydrolysis product) are formed. On the other hand, time course monitoring of the reaction of E383A + 200 mM azide reveals the formation of a doublet at $\delta$ 5.55 ppm ($J =$ 3.6 Hz), assigned to H-1 of the α-glucosyl azide product, which parallels the disappearance of the doublet at $\delta$ 5.45 ppm ($J = 7.5$ Hz) of the 2,4-DNPG substrate. Both reactions are complete, and the azide adducts formed are stable after long incubation times. On the basis of the rationale of the chemical rescue methodology, these results provide functional evidence of Glu178 as being the general acid/base residue and Glu383 the catalytic nucleophile.

Kinetics of enzyme reactivation for the E178A, E178Q, and E383A mutants are summarized in Table 4. For the

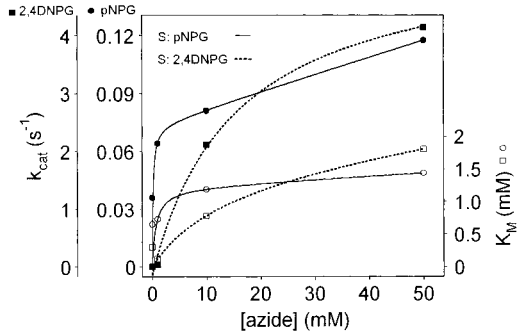



FIGURE 6: Kinetics of chemical rescue of the E178A mutant by sodium azide with pNPG (solid line) and 2,4-DNPG (dotted line) substrates. Solid symbols are for $k_{cat}$ values and open symbols for $K_M$ values.

E178A mutant with pNPG substrate, just a 3-fold increase in $k_{cat}$ is observed upon addition of 50 mM sodium azide, whereas a 200-fold increase is obtained with the more activated 2,4-DNPG substrate. As shown in Figure 6, $K_M$ values also follow an increase with azide concentration that parallels the effect on $k_{cat}$, thus keeping $k_{cat}/K_M$ approximately constant. Since $k_{cat}/K_M$ reflect the first irreversible step, most probably the glycosydation step to form the glycosyl-enzyme intermediate, its constant value across the range of azide concentrations studied is consistent with the effect of the azide anion as a nucleophile primarily on the deglycosylation step and not on glycosylation. Likewise, the increase in $K_M$ values with increasing azide is a consequence of the relative decrease of the amount of accumulated intermediate as the deglycosylation step speeds up.

The kinetic behavior of chemical rescue obtained for the Bgl3 enzyme is similar to that reported for the exoglucanase/xylanase Cex from *Cellulomonas fimi* (*37*): the activity of the Ala mutant at the general acid/base residue with an activated substrate (2,4-dinitrophenyl $\beta$-glycoside) was restored by addition of azide and also, although to a lesser extent, with a less activated *p*-nitrophenyl glycoside substrate, consistent with *deglycosylation* being rate determining for both substrates. For the Abg $\beta$-glucosidase from *Agrobacterium* (*11*), the corresponding mutant E170G is reactivated with the 2,4-DNPG substrate, but not with pNPG, for which *glycosylation* was rate limiting, whereas rate-determining *deglycosylation* operated on the 2,4-DNPG substrate. The low but significant reactivation of the Bgl3 E178A mutant here reported with pNPG indicates that for this substrate the *deglycosylation* step is still partially rate determining, and it becomes fully rate determining for the more activated 2,4-DNPG where a larger reactivation by azide is obtained. Regardless of the small differences, these systems have in common that azide operates on the *deglycosylation* step since $k_{cat}/K_M$ values are kept approximately constant on azide concentration. However, it seems not to be general for other retaining glycosidases as shown, i.e., for the *B. licheniformis* 1,3-1,4-$\beta$-glucanase (*29*) where $k_{cat}$ increases (activation) but $K_M$ values decrease with azide concentration, indicating that the exogenous nucleophile not only reactivates the mutant enzyme at the general acid/base residue as nucleophile in the *deglycosylation* step but also has an effect on the *glycosylation* step.

The nucleophile-less E383A mutant of Bgl3 shows no significant reactivation by azide with pNPG substrate (Table

4), indicating that the *p*-nitrophenol aglycon is not a good enough leaving group for the single displacement reaction to give the α-glucosyl azide product. With the more activated 2,4-DNPG substrate, the steady-state activity is partially restored, with a 100-fold increase in $k_{cat}$ and also an increase in $K_M$ at 1 M azide concentration.

## CONCLUSIONS

As a member of glycosyl hydrolases family 1, the *β*-glucosidase Bgl3 from *Streptomyces* shares similar mechanistic and kinetic properties with other family 1 *β*-glucosidases either from mesophiles [i.e., Agb from *Agrobacterium* (*13*)] or from thermophiles [i.e., Bgl from *Pyrococcus* (*18*)]. Substrate specificity, however, reveals that the Bgl3 enzyme is highly active on fucosyl glycosides, for which no substrate inhibition is observed as opposed to the glucosyl, galactosyl, and xylosyl glycoside substrates. Reactivity studies on the wt enzyme with aryl *β*-glucosides indicate that the glycosylation step is rate determining for the less reactive substrate (p$K_a$ of the phenol leaving group > 8) whereas deglycosylation becomes rate determining for more reactive substrates.

The identity of the catalytic residues, initially inferred from sequence alignments with other well-known *β*-glucosidases, has been confirmed by site-directed mutagenesis (Glu to Gln and Ala mutations) and their role in catalysis assessed by chemical rescue of the inactive alanine mutants at each residue with addition of azide as the exogenous nucleophile. In conclusion, Glu178 is the general acid/base residue, and Glu383 acts as the catalytic nucleophile in the retaining mechanims of Bgl3. Kinetics of mutant (E178A) reactivation by azide reveal subtle differences with other glycosidases, which together with the different substrate inhibition shown by the wt enzyme, indicate that there are subtle differences in the fine-tuning of the enzyme activity but that family 1 *β*-glucosidases share similar catalytic properties.

## REFERENCES


1. Perez-Pons, J. A., Cayetano, A., Rebordosa, X., Lloberas, J., Guasch, A., and Querol, E. (1994) *Eur. J. Biochem. 223*, 557−565.
2. Montero, E. (1998) Ph.D. Thesis, Universidad Complutense de Madrid.
3. Montero, E., Vallmitjana, M., Pérez-Pons, J. A., Querol, E., Jiménez-Barbero, J., and Cañada, F. J. (1998) *FEBS Lett. 421*, 243−248.
4. Guasch, A., Vallmitjana, M., Pérez, R., Querol, E., Pérez-Pons, J. A., and Coll, M. (1999) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 679−682.
5. Sinnott, M. L. (1990) *Chem. Rev. 90*, 1171−1202.
6. Davies, G., Sinnott, M. L., and Withers, S. G. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M. L., Ed.) pp 119−209, Academic Press Ltd., London.
7. Heightman, T. D., and Vasella, A. (1999) *Angew. Chem., Int. Ed. Engl. 38*, 750−770.
8. Henrissat, B., and Bairoch, A. (1993) *Biochem. J. 293*, 781−788.
9. McCarter, J. D., and Withers, S. G. (1994) *Curr. Opin. Struct. Biol. 4*, 885−892.
10. Trimbur, D. E., Warren, R. A. J., and Withers, S. G. (1992) *J. Biol. Chem. 267*, 10248−10251.
11. Wang, Q., Trimbur, D. E., Graham, R., Warren, R. A. J., and Withers, S. G. (1995) *Biochemistry 34*, 14554−14562.
12. Withers, S. G., Warren, R. A. J., Street, I. P., Rupitz, K., Kempton, J. B., and Aebersold, R. (1990) *J. Am. Chem. Soc. 112*, 5887−5889.
13. Kempton, J. B., and Withers, S. G. (1992) *Biochemistry 31*, 9961−9969.
14. Street, I. P., Kempton, J. B., and Withers, S. G. (1992) *Biochemistry 31*, 9970−9978.
15. Namchuk, M. N., and Withers, S. G. (1995) *Biochemistry 34*, 16194−16202.
16. Withers, S. G. (1995) in *Carbohydrate Bioengineering* (Petersen, S. B., Svensson, B., and Pedersen, S., Eds.) pp 97−111, Elsevier, Amsterdam.
17. Koshland, D. E., Jr. (1953) *Biol. Rev. 28*, 416−436.
18. Bauer, M. W., and Kelly, R. M. (1998) *Biochemistry 37*, 17170−17178.
19. Perez-Pons, J. A., Padros, E., and Querol, E. (1995) *Biochem. J. 308*, 791−794.
20. Dess, D., Kleine, H. P., Weinberg, D. V., Kaufman, R. J., and Sidhu, R. S. (1981) *Synthesis*, 883−885.
21. Malet, C., Viladot, J. L., Ochoa, A., Gállego, B., Brosa, C., and Planas, A. (1995) *Carbohydr. Res. 274*, 285−301.
22. Sharma, S. K., Corrales, G., and Penades, S. (1995) *Tetrahedron Lett. 36*, 5627−5630.
23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
24. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
25. Gill, S. C., and van Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
26. Juncosa, M., Pons, J., Dot, T., Querol, E., and Planas, A. (1994) *J. Biol. Chem. 269*, 14530−14535.
27. Clarke, J., and Fersht, A. R. (1993) *Biochemistry 32*, 4322−4329.
28. Pons, J., Planas, A., and Querol, E. (1995) *Protein Eng. 8*, 939−945.
29. Viladot, J. L., de Ramon, E., Durany, O., and Planas, A. (1998) *Biochemistry 37*, 11332−11342.
30. Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J., and Hughes, M. A. (1995) *Structure 3*, 951−960.
31. Painbeni, E., Valles, S., Polaina, J., and Flors, A. (1992) *J. Bacteriol. 174*, 3087−3091.
32. Gabelsberger, J., Liebl, W., and Schleifer, K.-H. (1993) *FEMS Microbiol. Lett. 109*, 131−138.
33. Nucci, R., Moracci, M., Vaccaro, C., Vespa, N., and Rossi, M. (1993) *Biotechnol. Appl. Biochem. 17*, 239−250.
34. Fersht, A. R. (1988) *Biochemistry 27*, 1577−1580.
35. Wang, Q., Graham, R. W., Trimbur, D. E., Warren, R. A. J., and Withers, S. G. (1994) *J. Am. Chem. Soc. 116*, 11594−11595.
36. Moracci, M., Trincone, A., Perugino, G., Ciaramella, M., and Rossi, M. (1998) *Biochemistry 37*, 17262−17270.
37. Macleod, A. M., Lindhorst, T., Withers, S. G., and Warren, R. A. J. (1994) *Biochemistry 33*, 6371−6376.
38. Planas, A. (1998) in *Carbohydrases from Trichoderma reesei and other microorganisms* (Claeyssens, M., Nerinckx, W., and Piens, K., Eds.) pp 21−38, The Royal Society of Chemistry, Cambridge.
39. Richard, J. P., Huber, R. E., Lin, S., Heo, C., and Amyes, T. L. (1996) *Biochemistry 35*, 12377−12386.
40. Moracci, M., Capalbo, L., Ciaramella, M., and Rossi, M. (1996) *Protein Eng. 9*, 1191−1195.


BI002947J

# II.3. ARTICLE TERCER

**Especificitat de substrat de la β-glucosidasa Bgl3: anàlisi del paper de la cisteina 181**

## II.2.a. ESPECIFICITAT DE SUBSTRAT DE LA β-GLUCOSIDASA BGL3: ANÀLISI DEL PAPER DE LA CISTEINA 181

### Substrate specificity in a family 1 β-glucosidase: analysis by site directed mutagenesis of the role of a cysteine residue located in the active site

En aquest article s'exposa l'anàlisi cinètica realitzada sobre dotze substitucions diferents a la cisteïna-181 de la β-glucosidasa Bgl3 d'*Streptomyces* sp. La Cys-181 es troba relativament ben conservada entre les β-glucosidases bacterianes de la família 1 i s'ha observat per anàlisi estructural que està localitzada en el centre actiu a la zona d'influència del subseti +1. La família 1 de les glicosilhidrolases presenta una baixa especificitat per la meitat reductora del substrat; no obstant això, estudiant la seva activitat sintètica per transglicosidació o per l'estratègia de la glicosintasa, s'ha observat que l'enzim prefereix determinades configuracions a l'aglicó. Les interaccions de l'enzim amb el substrat al subseti +1 són, però, difícils de determinar: no s'ha trobat un inhibidor que mostri les interaccions de l'enzim sobre la meitat reductora del substrat. Els estudis indirectes amb substrats amb substitucions als seus hidroxils indiquen que les interaccions sobre aquesta meitat del substrat són més febles que amb l'extrem no reductor.

S'ha substituït la posició C181 pels aminoàcids polars Ser i Gln, els hidrofòbics Ala, Val, Pro, i Leu, els hidrofòbics aromàtics Phe, Tyr, Trp, els carregats positivament Arg, Lys i els carregats negativament Asp i Glu. El mutant C181W s'ha expressat totalment insoluble i no s'ha pogut caracteritzar. S'ha comprovat el correcte plegament de tots els mutants per seguiment de la desnaturalització amb urea. L'anàlisi cinètica s'ha fet amb dos substrats: cel·lobiosa i p–nitrofenil-glucòsid (pNPG). La variació en l'activitat per als diferents mutants és més elevada amb cel·lobiosa ($k_{cat}$ de 170 $s^{-1}$ a 0,001 $s^{-1}$ i $k_{cat}/K_M$ de 8,6 $s^{-1}mM^{-1}$ a $1{,}1\cdot10^{-6}$ $s^{-1}mM^{-1}$) que amb pNPG ($k_{cat}$ de 110 $s^{-1}$ a 0.11 $s^{-1}$ i $k_{cat}/K_M$ de 310 a 0,11 $s^{-1}mM^{-1}$).

Les substitucions per un residu hidrofòbic rendeixen enzims amb una $K_M$ molt reduïda respecte al wt per al substrat pNPG. Amb cel·lobiosa la $K_M$ és molt depenent del volum del residu hidrofòbic que substitueix la cisteïna; així per a hidrofòbics petits la $K_M$ es troba al voltant de 10-20 vegades el wt, i pels aromàtics no s'observa saturació en el marge de concentracions de substrat possible (fins a 360 mM cel·lobiosa).

Les substitucions pels residus polars, glutamina i serina, dóna enzims similars al wt en termes d'activitat pels dos substrats provats, però amb valors de $K_M$ més elevats. De fet el mutant C181S sembla comportar-se més com si fos un residu hidrofòbic petit que no pas un de polar; una simulació estructural del mutant mostra com l'hidroxil de la serina s'enfonsa per interaccionar amb la cadena alfa de la proteïna, orientant a la cavitat del centre actiu la part hidrofòbica de la seva cadena lateral.

Les substitucions per aminoàcids carregats produeixen resultats aparentment incoherents: les parelles de mutants C181R i C181K amb càrrega positiva, i els acídics C181D i C181E tenen un comportament diferent entre ells quan s'esperaria una certa homogeneïtat. Amb pNPG C181K és 50 vegades més inactiu que C181R; de manera similar C181D és 100 vegades més inactiu que C181E, essent la $K_M$ també 10 vegades més alta per C181D que per C181E. Amb cel·lobiosa com a substrat s'obtenen resultats semblants però en conjunt es pot dir que les substitucions donen lloc a enzims molt inactius. Les diferències es poden atribuir a la diferent longitud de la cadena lateral de cada residu i, per tant, petites variacions en la seva orientació o arquitectura fina en el centre actiu explicarien el comportament dels mutants. La Cys-181 es troba molt propera al Glu-178, residu àcid/base general de la catàlisi; així, les diferències observades entre aquests mutants s'explicarien probablement per l'efecte sobre el p$K_a$ del Glu-178.

En general, doncs, i en termes d'eficiència catalítica ($k_{cat}/K_M$) els mutants per residus amb caràcter hidrofòbic conserven millor les propietats de l'enzim salvatge. Entre aquests, cal destacar el mutant C181V; curiosament, en nombroses glucosidases de la família 1 es troba un residu valina en la posició de la Cys-181 a Bgl3. Les dades obtingudes semblarien indicar que el residu cisteïna, amb un caràcter intermedi entre hidrofòbic i polar, és el residu "ideal" per aquesta posició. En aquest sentit, es podria considerar a la Bgl3 (i per extensió a les altres glucosidases de la família) com enzims molt evolucionats o, fins i tot, molt perfeccionats.

A la vista dels resultats, caldria atribuir a la Cys-181 una funció d'assistència a la correcta orientació del substrat en el centre actiu, a través de la interacció de manera principalment hidrofòbica amb la meitat reductora d'aquell (probablement sobre el centre de l'anell).

# Substrate specificity in a family 1 β-glucosidase: analysis by site directed mutagenesis of the role of a cysteine residue located in the active site

**Miquel Vallmitjana, Raquel Planell, Mario Ferrer, Enrique Querol i Josep A. Pérez-Pons**

Institut de Biotecnologia i Biomedicina "Vicent Villar i Palasí" i Dept. de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona (Spain)

## Abstract

In order to elucidate the role of the cysteine residue at position 181 in the β-glucosidase Bgl3 (EC 3.2.1.21) from *Streptomyces* sp. (ATCC 11238), thirteen site-directed mutants were constructed. Such a residue is relatively well conserved between family 1 bacterial β-glucosidases and belongs to the subsite +1 in the more external side of the pocket conforming the active centre. Residues involved in subsite +1 would interact with the aglycon moiety of the sugar thus determining the substrate-specificity differences among β-glucosidases. Family 1 enzymes show however low specificity for the aglycon and no precise data on the nature of such interactions are available since a high disorder has been observed in most of the inhibitor or substrate analogue atoms in the tridimensional complex structures solved so far. Kinetic studies using two model substrates, p-nitrophenyl-glucoside (pNPG) and the disaccharide cellobiose, indicate that this position is closely associated with the enzyme activity and substrate specificity. With the aryl-glucoside pNPG replacement of Cys181 by hydrophobic amino acids greatly decreases $K_M$ values respect to the wild type enzyme while with cellobiose $K_M$ is dependent of the hydrophobic side-chain volume and even no substrate saturation (at least to 360 mM cellobiose) was observed in the case of aromatic substitutions. In general, the more inactive mutants correspond to replacements introducing a charged residue. Mutants showing $k_{cat}$ values closer to the wild-type for both substrates correspond to replacements by polar and small hydrophobic residues. Regarding the catalytic efficiency (in terms of $k_{cat}/K_M$), hydrophobic mutants (particularly, L, F, and Y) show a large decrease in cellobiase activity while maintain the pNPGase activity at levels close to the wild-type enzyme. On the contrary, polar replacements (S and Q) conserve the catalytic efficiency on cellobiose but no on pNPG. These results can be interpreted as a shift in substrate specificity or preference.

## Introduction

β-Glucosidases (β-glucoside glucohydrolases, E.C. 3.2.1.21) are enzymes which act on disaccharides, oligosaccharides or conjugated glycosides catalysing the hydrolysis of the β-glycosidic bond by means of a double displacement mechanism [1,2] with formation of a

covalent glycosyl-enzyme intermediate (*glycosylation* step) and subsequent release of the free sugar (*deglycosylation* step). As the CAZY database (http://afmb.cnrs-mrs.fr/CAZY) shows, β-glucosidases are classified into families 1 and 3 into the large group of glycosyl hydrolases according to amino acid sequence similarity and hydrophobic cluster analysis. Such a classification also evidences the conservation of the catalytic machinery and a common fold for members of a same family. In particular family 1 includes more than 280 proteins from archaebacteria to humans and it is characterized both, by a retention mechanism in which usually two carboxylate (glutamic acid) residues are involved, and by an $(\alpha/\beta)_8$-barrel structural fold. However, in spite of the large amount of sequences (and some 3D-structures in Protein Data Bank), there are very few works dealing with the protein determinants of substrate recognition and specificity which is an important aspect for example for biotechnological applications [3]. Two subsites have been defined In the active centre of these enzymes: subsite -1 which accomodates the glycone moiety or non-reducing end of the substrate, and subsite +1 which will interact with the aglycon moiety or reducing end [4]. The lack of knowledge is particularly true for subsite +1 where would lie the basis for the multiplicity of "preferred" enzymatic activities shown by β-glycosidases (i.e., aryl-glycosidases, alkyl-glycosidases, cellobiases, oligosaccharidases, etc). Thus, while residues involved in recognition of the glycone moiety seem to be already defined from several structures of substrate/inhibitor-enzyme complexes [5,6,7], those interacting with the aglycon side are not so well characterized mainly because of the high disorder observed for this side in crystals and a lesser sequence conservation in this protein region.

In the present study we have investigated the possible role of Cysteine-181 in the β-glucosidase Bgl3 from *Streptomyces* sp. (ATCC 11238). The enzyme Bgl3 (GenBank accession code Z29625) is a member of family 1 of glycosyl hydrolases we have functionally characterized [8] and recently solved its three-dimensional structure (PDB accession code 1GNX) complexed with the substrate analogues thiocellobiose and tetrathiocellobiose (manuscript in preparation). The Cys-181 is located at subsite +1 of the active centre and upon alignment of sequences belonging to family 1 glycosidases shows a remarkable degree of conservation specially among enzymes of prokaryotic origin. In particular, considering the set of bacterial family 1 members, about a 50% of the sequences contains a cysteine residue at the equivalent position being the alternative residue mostly a valine. Cysteine or valine residues are also frequently found in archaebacterial and eukaryotic β-glycosidases. Figure 1 shows a partial sequence alignment of the region comprising the β-sheet$_4$ and the N-terminal end of the α-helix$_4$ where Bgl3 Cys-181 is located; in the figure only sequences belonging to enzymes well characterized biochemically or with available structural data are included.

The role of Cys-181 residue has been studied by preparing a series of Bgl3 mutants at position 181; such mutant isoforms were kinetically analyzed using p-nitrophenyl-β-D-glucoside and cellobiose as model substrates for aryl-glucosidase and disaccharidase activities, respectively.

**Materials and methods**

*Protein Obtention.* Bgl3 enzymes were expressed and purified from recombinant *E. coli* BL21(DE3) strains carrying derivatives of pET-21d vector (Novagen) as previously described[8]. The fusion of a His-tag to the N-terminal end of Bgl3 allowed the rapid protein purification by $Ni^{2+}$-affinity chromatography. Protein concentration was determined by the dye-binding method of Bradford (1974) [9] using BSA as a standard.

*Site-Directed Mutagenesis*. The mutant C181A was obtained by PCR-mutagenesis according to the method reported by Juncosa et al. (1997) [10] and using as template a 560 bp *Sac* I-*Hin*d II fragment from the bgl3-gene, which includes Cys-181 residue. Briefly, in the first PCR run mutagenic primer C181A (Table 1) and an internal oligonucleotide were used as primers. The PCR product from the first run was purified by agarose gel electrophoresis and used as a primer (204 bp in length) for a second PCR run together with other internal oligonucleotide. The mutated fragment was purified, cloned in pUC118 vector, and completely sequenced before its replacement by the corresponding fragment in the recombinant plasmid containing the wild-type gene (namely, pET21-HBG3). In order to facilitate the obtention of the other Cys-mutants a "cassette" strategy was applied using the synthetic oligonucleotides indicated in Table 1 after the introduction by silent mutagenesis of two new restriction sites flanking the Cys-181 codon of the *bgl3* gene. DNA sequencing was done by means of the Big-dye kit for automatic DNA sequencers (Perkin-Elmer).

*Equilibrium Urea Denaturation*. Unfolding of wild type and mutant Bgl3 isoforms was monitored by measuring the dependence of fluorescence intensity on urea concentration in the range 0-8 M, as previously reported [8].

*Kinetic Characterization*. β-Glucosidase activity was measured using p-nityrophenyl-β-D-glucoside (pNPG) and cellobiose as substrates. *(A) p-Nitrophenyl-β-D-glucoside* hydrolysis was determined by a continuous assay monitoring the release of p-nitrophenol from pNPGlu at 400 nm in a thermostated spectrophotometer. Enzyme reactions were performed in 50 mM sodium phosphate buffer, pH 6.5, at 50 ºC using different pNPGlu concentrations (0.004-23 mM). In these conditions, the molar absorption coefficient at 400 nm was 5486 $M^{-1}cm^{-1}$. The protein concentrations in the reaction mixtures were 3-260 nM. *(B) Cellobiose* hydrolysis was determined by a two-steps assay as follows: (1) reaction mixtures containing different cellobiose (0.5-360 mM) and enzyme (0.01-10 μM) concentrations in 50 mM sodium phosphate buffer, pH 6.5, were incubated for 1-5 min at 50 ºC (for highly inactive mutants, incubation time was 5h and to preserve 100% activity during such long incubation times pH and temperature were changed to 7.0 and 30 ºC, respectively); (2) the glucose released from cellobiose was

measured using the glucose ultraviolet method (Boehringer Mannheim) after stopping the former reactions by heating at 95-100 ºC for 5 min.

Hydrolysis rates were obtained from the initial slopes and kinetic constants were calculated from data fitted to the Michaelis-Menten equation or to a substrate inhibition model ($v = k_{cat}[E]_0[S]/(K_M + [S] + [S]^2/K_i)$) by means of nonlinear regression analysis.

*Structural analysis*. Alingments were performed by the Clustal method (www.ebi.ac.uk/clustalw). The structure of the different enzymes mutated at position 181 was modelled on the three-dimensional structures corresponding to Bgl3 (PDB code 1GNX) and myrosinase-2-deoxy-2-fluoroglucoside complex (PDB code 1E70)[11]. Molecular modelling was carried out using the SwissProt PDB Viewer program and applying an energy minimization of four cycles of 400 steps. The differences between wt and mutant enzymes were evaluated by superimposing structures and comparing distances between selected atoms.

**Results**

In order to investigate the role of Cys-181, a total of thirteen different Bgl3 β-glucosidases were obtained by site-directed mutagenesis in which the wild-type cysteine residue was replaced by either three hydrogen-bonding (S, Q, Y), six non-hydrogen-bonding (A, V, L, P, F, W), and four charged (R, K, D, E) amino acids. The Bgl3 isoforms were expressed as mainly soluble proteins with yields similar to the wild-type enzyme. Nevertheless, it was not possible to analyse the C181W mutant since it was produced in a completely insoluble form and we were not able to purify even minute amounts of soluble protein, likely indicating strong folding/structural constraints for that substitution. During purification of the different mutated enzymes special cautions were taken in order to ensure the absence of contamination, particularly from wild-type, that would mask the activity assays altering results and conclusions. With this aim single His-bind columns (Pharmacia) were used for each enzyme and the chromatography system was throughly washed with 0.5 M NaOH before and after any purification process. The folding or stability of the mutants was evaluated by equilibrium urea denaturation monitoring the fluorescence intensity at urea concentrations from 0 to 8 M. Figure 2 shows the denaturation curves obtained for wild-type Bgl3 and representative mutants. Considering that $[D]_{50\%}$ is the concentration of denaturant at which 50% of the protein is unfolded, the mean $[D]_{50\%}$ value between mutant enzymes was 3.04 M urea (with a standard deviation of ±0.15). Such a value compares well with $[D]_{50\%}$ calculated for the wild-type (i.e., 3.44 M) as do the denaturation profiles. Structures of the mutants were also checked by energy minimization modelling on the wild-type three-dimensional structure (PDB code 1GNX) and no gross changes or rearrangements in the active site of the enzyme were observed. These results support a folded structure for the mutant proteins, and hence, changes of the catalytic properties can be assigned to the replaced amino acid residue.

Kinetics for the wild-type and mutant Bgl3 glucosidases were evaluated with p-nitrophenyl-β-D-glucoside and the disaccharide cellobiose (β-1,4-glucosyl-glucoside), as summarized in Tables 2 and 3, respectively. The position 181 is located in the active site of the enzyme, specifically at the +1 subsite where interactions with the aglycon moiety of the substrate can be done (Figure 3). On the other hand, pNPG and cellobiose afford two aglycons with different characteristics and reactivities and for this reason were chosen as model substrates for the kinetic studies. Results obtained with pNPG show that replacement of Cys-181 by an hydrophobic non-hydrogen-bonding residue slightly decreases, in comparison with wild-type, $K_M$ values while hydrogen-bonding or charged replacements increase such a value about 10-fold. A linear relationship between the hydrophobic character (according to current hydrophobicity scales) and decrease in the $K_M$ value can also be observed except for C181F mutant in which, despite its highly hydrophobic character, the large side chain volume may lead to steric hindrances affecting the reaction, as indicated by modelling this mutant on the structure of myrosinase-2-deoxy-2-fluoroglucoside complex [5]. Other exceptions to these general trends would be C181E and C181Q: the former shows a $K_M$ value equal to wild-type and the later a 100-fold greater value. In contrast, using cellobiose, all substitutions increase $K_M$ values for this substrate from 10- to 300-folds, corresponding the larger increases to mutants with larger side chains independently of the chemical character of the substituting amino acid. Thus, the steric or volume factor would be important for the reaction with cellobiose, a substrate with a larger aglycon moiety compared to pNPG.

Regarding the catalytic constant ($k_{cat}$), in general all mutants show greatly decreased values for both tested substrates; only mutant C181Q presents a significant increase in $k_{cat}$ values, so its increased activity is likely due to the high $K_M$ value. Mutants showing $k_{cat}$ values closer to the wild-type correspond to replacements by polar and small hydrophobic residues. The more important reductions in $k_{cat}$ were obtained with substitutions incorporating a charged amino acid residue (except C181E). Plotting log of $k_{cat}$ for pNPG *vs.* log of $k_{cat}$ for cellobiose (Figure 4) shows a straight correlation between both substrates, i. e. to a decrease in the activity against pNPG corresponds a decrease in the activity against cellobiose. The analysis of the catalytic efficiency (in terms of $k_{cat}/K_M$) shows differences between pNPG and cellobiose although in general all mutations have a negative effect on $k_{cat}/K_M$ values (Tables 2 and 3); exceptions are C181P and C181V mutants with $k_{cat}/K_M$ for pNPG (expressed as % of the wild-type) of 180% and 110%, respectively. Using the more hydrophobic substrate (i. e., pNPG), mutants introducing a hydrophobic aliphatic side chain (A, V, L) show $k_{cat}/K_M$ values higher than those with an aromatic group (F, Y); C181F and C181Y display low $K_M$ values similar to aliphatic substitutions but the activity in terms of $k_{cat}$ is lower yielding in turn a lower $k_{cat}/K_M$ value. In contrast, with the more polar cellobiose, the higher $k_{cat}/K_M$ values correspond to polar substitutions (S, Q) while among hydrophobic mutants a clear inverse relationship between catalytic efficiency and side-chain volume can be established. Mutations introducing a charge (R, K, D) also show a very low $k_{cat}/K_M$; instead C181E, being its $K_M$ extremely high (i. e., $10^3$-fold

the wild-type), keeps a relatively high catalytic efficiency. These observations are illustrated in Figure 5 by plotting log $k_{cat}/K_M$ (pNPG) *vs.* log $k_{cat}/K_M$ (cellobiose) values of the different Bgl3 glucosidases.

An interesting effect observed in several Cys-181 mutants is the loss of the excess-substrate inhibition shown by the wild-type enzyme with different aryl-glucosides as substrates but no with cellobiose or other β-linked disaccharides [8]. Among the mutants analyzed in the present work, only C181V and C181D show such inhibition with calculated $K_i$ values of 1.5 and 82.3 mM, respectively (1.4 mM for the wild-type). Inhibition by an excess of substrate is rather common in β-glucosidases [11,12] and, in turn, difficult to explain. However, a working hypothesis can be drawn considering the binding of a second substrate molecule to the glycosyl-enzyme intermediate (see below); this second molecule would hamper the departure of the reaction product. Moreover, the Bgl3 glucosidase also performs very efficiently transglycosylation reactions [13,14] and, in this sense, the synthetic reaction may mask the hydrolytic activity during the assays at "high" substrate concentration [15].

**Discussion**

As a retaining glycosidase, the mechanism of Bgl3 involves a two-step process [1,2]. In the first step (*glycosylation*) the amino acid residue acting as a general acid protonates the glycosidic oxygen with concomitant C-O breaking of the scissile glycosidic bond, while the deprotonated carboxylate functioning as a nucleophile attacks the anomeric center to give a covalent glycosyl-enzyme intermediate (E-G). The second step (*deglycosylation*) involves the attack of a water molecule assisted by the conjugate base of the general acid to render the free sugar with overall retention of configuration (G-OH).

$$E + G\text{-}X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\cdot G\text{-}X \xrightarrow{k_2} E\text{-}G \xrightarrow[H_2O]{k_3} E + G\text{-}OH$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad X$$

In this mechanism, $k_{cat} = k_2 \cdot k_3/(k_2+k_3)$, $K_M = k_3 \cdot (k_2+k_{-1})/k_1 \cdot (k_3+k_2)$, and $k_{cat}/K_M = k_1 \cdot k_2/(k_2+k_{-1})$.

Among the thirteen amino acid residues analyzed in the present study at position 181 of the Bgl3 β-glucosidase and considering both assayed substrates, the wild-type cysteine dispalys the best kinetic behaviour. Hydrophobic substitutions decreases $K_M$ values for pNPG likely indicating interactions of the substrate with residue at position 181. This would be reinforced by the fact that, with a more polar aglycon moiety (as in cellobiose), $K_M$ values are better for polar (S, Q) or small hydrophobic (A, V, P) substitutions. In the case of C181S, interactions with the substrate are not so clear since, upon the modelled structure of the mutant protein, the side-chain of serine residue establishes hydrogen-bonding with the protein main-chain (not shown). The introduction of a charged residue has in general a deletereous effect on kinetic constants. As shown in Figure 3, position 181 is very close to Glu-178, the residue acting as general acid/base catalyst which needs a fine regulation of its p$K_a$ with kinetic values of 7.5 for the free

enzyme and 7.1 for the E-S complex [8] to assist properly in the catalysis: a new charge close to the catalytic residues could change the p$K_a$ seriously impairing the enzyme activity. The reactivity of the pNPG as a substrate allows its hydrolysis (i. e., departure of the leaving group) during the *glycosylation* step without acidic assistance while basidic assistance for the *deglycosylation* step is independent of the leaving group as seems to indicate that $k_{cat}$ and $k_{cat}/K_M$ values for pNPG of mutant C181D are, respectively, only 3- and 2-folds greater than those corresponding to the E178A mutant [8]. On the contrary, the limiting step for cellobiose hydrolysis is glycosylation, thus explaining the dramatic reduction of enzyme activity observed in positive-charge substituted mutants (R, K). The "anomalous" behaviour of the C181E, maintaining a reasonable activity in terms of $k_{cat}$ on both substrates, may be explained considering that the length of the residue side-chain can allow its separation from E178 which could, then, preserve the functional p$K_a$ values.

Since $k_{cat}/K_M$ reflects the first irreversible step, most probably the *glycosylation* step to form the glycosyl-enzyme intermediate, the results here presented may be analyzed in terms of substrate specificity. In this sense, some of the substitutions (e. g., L, F, Y) give rise to a remarkable shift in substrate preference: they lose cellobiase activity (in terms of $k_{cat}/K_M$) while pNPGase activity is kept at values close to the wild-type. In the opposite, Ser and Gln substitutions show an important reduction in catalytic efficiency against pNPG, maintaining almost constant respect to the wild-type the efficiency on cellobiose (Figure 5). Thus, results obtained with the mentioned mutants can be interpreted as a change or shift in their substrate specificity.

Cysteine is a flexible residue that could play a variety of roles, behaving in proteins similarly to tryptophan and methionine, i. e. acting with a character hydrophobic rather than polar [16]. Moreover, cysteine is structurally close to serine but could establish different kind of interactions, for instance abolishing hydrogen-bonding. Our results suggest that Cys-181 does not interact with substrate in a stringent specific way but it contributes to the orientation of the sugar molecule or, at least, guides its entry into the slot of the active site. Kinetic results also indicate that Cys in position 181 represents a very good solution affording both a high catalytic efficiency and a broad range of hydrolyzable substrates. Early criteria to distinguish between aryl-glucosidases and cellobiases include the differential inhibiton by p-chloromercuribenzoate, a thiol-specific reagent that selectively inhibited cellobiase but no aryl-glucosidase activity [17], suggesting therefore the existence of a cysteine residue in the region conforming the active site of some glucosidases. At the moment, assessing the determinants for the recognition and interaction of the enzyme with the substrate aglycon moiety has proven to be difficult by means of resolving the 3D-structures of enzyme-inhibitor or substrate analogue complexes due to the high disorder shown by inhibitor aglycon atoms. Nevertheless, it seems that the protein interacts tightly with the extreme end of the sugar ring, i. e. at the glycon moiety, through the formation of hydrogen bonds while the aglycon is probably hold by $\pi$-stacking interactions [7,18]. The utilization of site-directed mutagenesis combined with the resolution of the structure of enzyme-substrate complexes should provide definitive insights on the determinants of substrate

recognition and binding. These studies will also provide clues, or even enzymic tools (i. e., mutant proteins with improved capabilities), for tailoring interesting applications of glycosidases such as the synthesis of sugars.

## References


1 Sinnott, M. L. (1990) Catalytic mechanisms of enzymic glycosyl transfer. Chem. Rev. **90**, 1171-1202

2 Davies, G. J., Sinnott, M. L. and Withers, S. G. (1998) Glycosyl transfer. In Comprehensive Biological Catalysis: a mechanistic reference (Sinnott, M. L., ed.), vol 1, pp. 119-209, Academic Press, London

3 Ly, H. D. and Withers, S. G. (1999) Mutagenesis of glycosidases. Annu. Rev. Biochem. **68**, 487-522

4 Fernández, P., Cañada, F. J., Jiménez-Barbero, J. and Martín-Lomas, M. (1995) Substrate specificity of small-intestinal lactase: Study of the steric effects and hydrogen bonds involved in enzyme-substrate interaction. Carbohydr. Res. **271**, 31-42

5 Burmeister, W. P., Cottaz, S., Driguez, H., Lori, R., Palmieri, S. and Henrissat, B. (1997) The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. Structure **5**, 663-675

6 Wiesmann, C., Hengstenberg, W. and Schulz, G. E. (1997) Crystal structures and mechanism of 6-phospho-β-galactosidase from *Lactococcus lactis*. J. Mol. Biol. **269**, 851-860

7 Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B. and Esen, A. (2001) Crystal structure of a monocotyledon (maize ZMGlu1) β-glucosidase and a model of its complex with p-nitrophenyl-β-D-thioglucoside. Biochem. J. **354**, 37-46

8 Vallmitjana, M., Ferrer-Navarro, M., Planell, R., Abel, M., Ausín, C., Querol, E., Planas, A. and Pérez-Pons, J. A. (2001) Mechanism of the family 1 β-glucosidase from *Streptomyces* sp: catalytic residues and kinetic studies. Biochemistry **40**, 5975-5982

9 Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248-254

10 Juncosa, M., Pons, J., Dot, T., Querol, E. and 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

11 Patchett, M. L., Daniel, R. M. and Morgan, H. W. (1987) Purification and properties of a stable β-glucosidase from an extremely thermophilic anaerobic bacterium. Biochem. J. **243**, 779-787

12 Lin, J., Pillay, B. and Singh, S. (1999) Purification and biochemical characteristics of beta-D-glucosidase from a thermophilic fungus, *Thermomyces lanuginosus*-SSBP. *Biotechnol. Appl. Biochem.* **30**, 81-87

13 Pérez-Pons, J. A., Cayetano, A., Rebordosa, X., Lloberas, J., Guasch, A. and Querol, E. (1994) A β-glucosidase gene (bgl3) from *Streptomyces* sp. strain QM-B814. Molecular cloning, nucleotide sequence, purification and characterization of the encoded enzyme, a new member of family 1 glycosyl hydrolases. Eur. J. Biochem. **223**, 557-565

14 Montero, E., Alonso, J., Cañada, F. J., Fernández-Mayoralas, A. and Martín-Lomas, M. (1997) Regioselectivity of the enzymatic transgalactosidation of D- and L- xylose catalysed by β-galactosidases. Carbohydr. Res. **305**, 383-391

15 Harhangi, H. R., Steenbakkers, P. J. M., Akhmanova, A., Jetten, M. S. M., van der Drift, C. and Op den Camp, H. J. M. (2002) A highly expressed family 1 L-glucosidase with transglycosylation capacity from the anaerobic fungus *Piromyces* sp. E2. Biochim. Biophys. Acta **1574**, 293-303

16 Nagano, N., Ota, M. and Nishikawa, K. (1999) Strong hydrophobic nature of cysteine residues in proteins. FEBS Lett. **458**, 69-71

17 Lusis, A. J. and Becker, R. R. (1973) The β-glucosidase system of the thermophilic fungus *Chaetomium thermophile* var. coprophile. Biochim. Biophys. Acta **329**, 5-16

18 Marques, A. R., Coutinho, P. M., Videira, P., Fialho, A. M. and Sá-Correia, I. (2003) *Sphingomonas paucimobilis* β-glucosidase Bgl1: a member of a new bacterial subfamily in glycoside hydrolase family 1. Biochem. J. **370**, 793-804

**Table 1**. Oligonucleotide employed for the site-directed mutagenesis of Cys 181 residue. The mutagenic nucleotides are shown in lowercase.

| Mutant | Mutagenic Oligonucleotide Sequence |
|---|---|
| C181Y | TGCACGCCGGAGCCGTAGCCtAGGAACGCGCTgtaCCAtGGCT |
| | ATGGtacAGCGCGTTCCTaGGCTACGGCTCCGGCG |
| C181E | CTAGGAACGCGCTttcC |
| | CATGGgaaAGCGCGTTC |
| C181K | CTAGGAACGCGCTtttC |
| | CATGGaaaAGCGCGTTC |
| C181R | CTAGGAACGCGCTacgC |
| | CATGGcgtAGCGCGTTC |
| C181S | CTAGGAACGCGCTggaC |
| | CATGGtccAGCGCGTTC |
| C181V | CTAGGAACGCGCTaacC |
| | CATGGgttAGCGCGTTC |
| C181L | CTAGGAACGCGCTcagC |
| | CATGGctgAGCGCGTTC |
| C181F | CTAGGAACGCGCTgaaC |
| | CATGGttcAGCGCGTTC |
| C181W | CTAGGAACGCGCTccaC |
| | CATGGtggAGCGCGTTC |
| C181D | CTAGGAACGCGCTctgC |
| | CATGGcagAGCGCGTTC |
| C181P | CTAGGAACGCGCTtggC |
| | CATGGccaAGCGCGTTC |
| C181A | AACGAGCCCTGGgctAGCGCGTTCCTC |

**Table 2**. Kinetic parameters of the different Bgl3 isoforms for p-nitrophenyl-β-glucoside substrate. The asterisk indicates enzymes showing excess-substrate inhibition.

| **mutant** | $k_{cat}$ ($s^{-1}$) | Relative $k_{cat}$ (%) | $K_M$ (mM) | $k_{cat}/K_M$ ($s^{-1}$ $mM^{-1}$) | Relative $k_{cat}/K_M$ (%) |
|---|---|---|---|---|---|
| wt* | 41.67 | 100 | 0.15 | 277.8 | 100 |
| C181S | 10.47 | 25 | 1.49 | 7.01 | 2.5 |
| C181Q | 113.4 | 270 | 110 | 1.03 | 0.37 |
| C181A | 12.83 | 31 | 0.098 | 130.95 | 47 |
| C181V* | 7.63 | 18 | 0.0244 | 312.6 | 110 |
| C181L | 3.33 | 8.0 | 0.0192 | 174.04 | 63 |
| C181P | 32.84 | 79 | 0.065 | 505.2 | 180 |
| C181F | 2.43 | 5.8 | 0.0824 | 29.53 | 11 |
| C181Y | 3.04 | 7.3 | 0.0905 | 33.64 | 12 |
| C181R | 5.02 | 12 | 1.47 | 3.401 | 1.2 |
| C181K | 0.184 | 0.44 | 1.101 | 0.167 | 0.06 |
| C181D* | 0.114 | 0.27 | 1.023 | 0.111 | 0.04 |
| C181E | 12.41 | 30 | 0.167 | 74.48 | 27 |

**Table 3**. Kinetic parameters of the different Bgl3 isoforms for cellobiose substrate. Asterisk indicates that the value was calculated as the slope $v_0/[E] = k_{cat}/K_M \cdot [S]$ at low [S] when no saturation was observed.

| mutant | $k_{cat}$ ($s^{-1}$) | Relative $k_{cat}$ (%) | $K_M$ (mM) | $k_{cat}/K_M$ ($s^{-1}mM^{-1}$) | Relative $k_{cat}/K_M$ (%) |
|---|---|---|---|---|---|
| wt | 35.55 | 100 | 4.13 | 8.597 | 100 |
| C181S | 26.93 | 75.8 | 47.17 | 0.571 | 6.64 |
| C181Q | 167.98 | 473 | 92.4 | 1.82 | 21.2 |
| C181A | 23.93 | 67.3 | 76.78 | 0.312 | 3.63 |
| C181V | 0.33 | 0.934 | 33.63 | 0.0099 | 0.115 |
| C181L | >0.0005 | 0.0014 | >360 | 1.6 E-6 * | 1.9 E-05 |
| C181P | 25.78 | 72.5 | 77.22 | 0.33 | 3.89 |
| C181F | >0.03 | 0.08 | >360 | 8·E-5 * | 9.3·E-4 |
| C181Y | >0.024 | 0.07 | >360 | 7.6 E-5 | 8.8 E-04 |
| C181R | >0.0006 | 0.0017 | >360 | 2.2·E-6 * | 2.6·E-05 |
| C181K | < 5E-4 | 1.4e-03 | - | - | - |
| C181D | 0.0027 | 0.0077 | 89.5 | 3.1·E-5 | 3.6·E-4 |
| C181E | > 3 | 8.4 | >350 | 0.009 * | 0.1 |

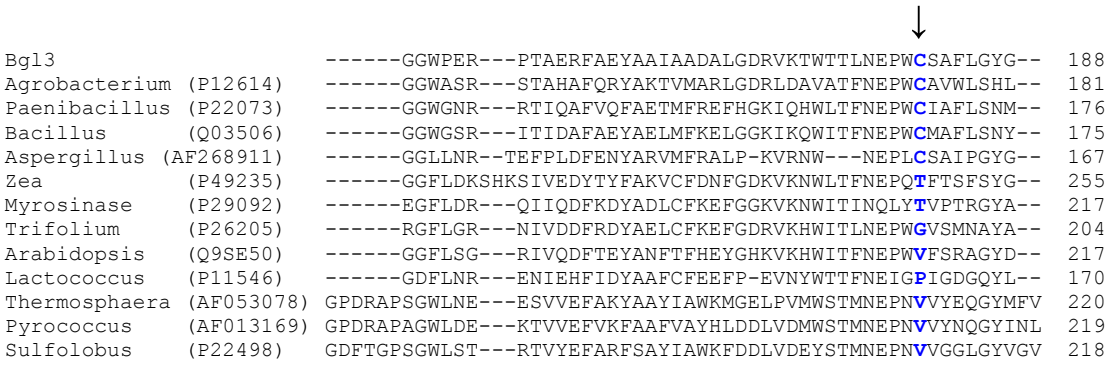



**Figure 1**. Partial sequence alingment of the region corresponding to the β4-sheet in family 1 β-glucosidases. SwissProt accession numbers are shown into parentheses. The arrow shows the Cysteine 181.

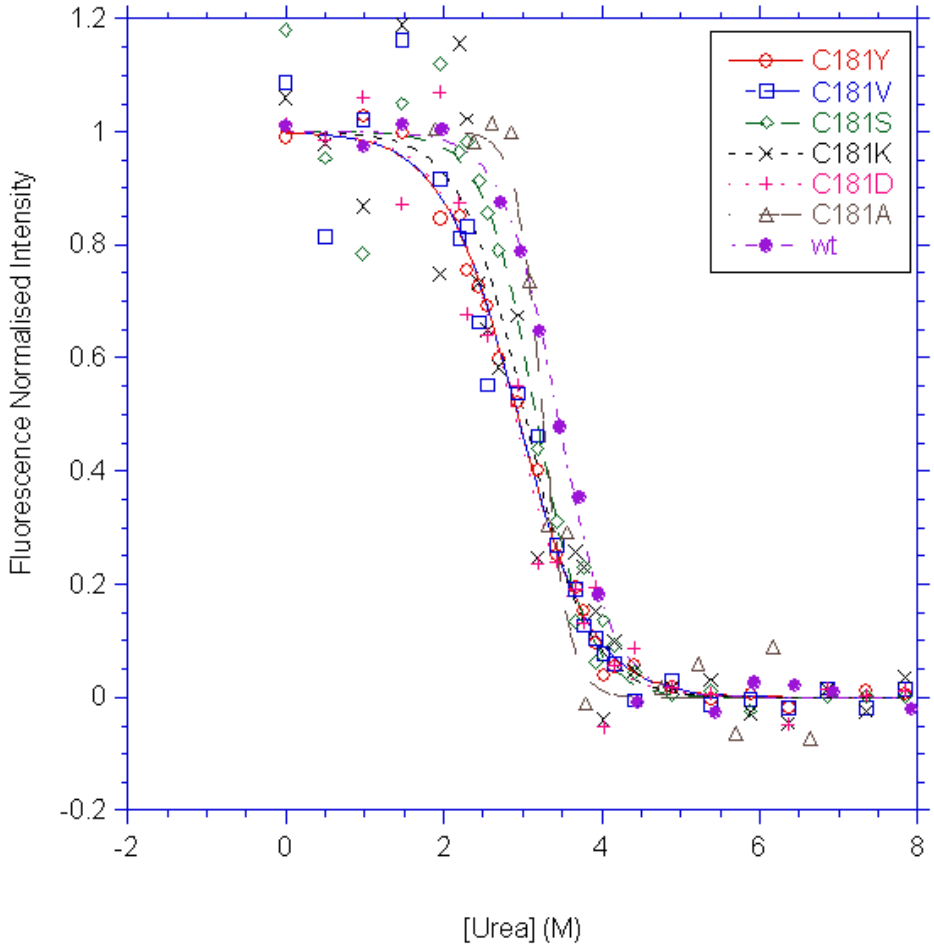



**Figure 2**. Equilibrium urea denaturation curves for the wild-type and mutant β-glucosidases. Fluorescence intensity values were fitted to $F = (a_F + b_F \cdot [D] + (a_U + b_U[D])\exp(m[D] - [D]_{50\%})/RT))(1 + \exp(m([D] - [D]_{50\%})/RT))$ (eq. 1). Normalized curves are plotted as the fraction native $f_N = (F - (a_U + b_U[D]))/(a_F - a_U + [D](b_F - b_U))$ (eq. 2), where $F$ is the measured fluorescence intensity and $a_U$, $a_F$, $b_U$, and $b_F$ are the adjusted parameters from eq. 1.

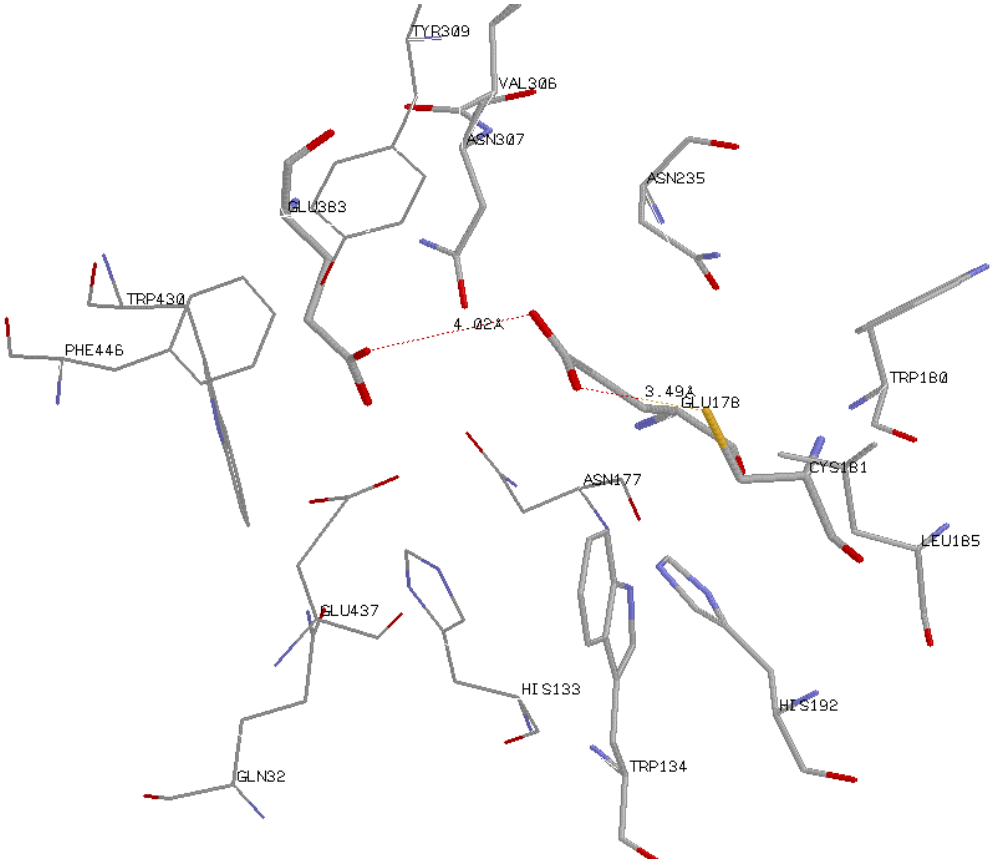



**Figure 3**. Detail of the active centre of Bgl3 β-glucosidase (PDB accession code 1GNX).

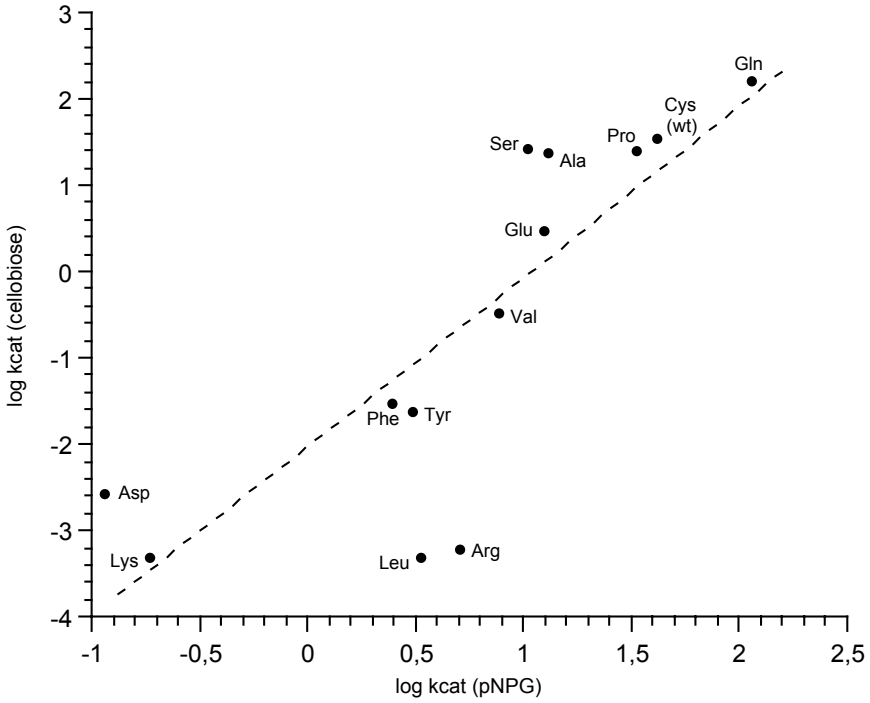



**Figure 4**. Catalytic rate. Log$k_{cat}$ relationship between cellobiose and p-nitrophenyl-β-D-glucoside substrates for Bgl3 mutants at position 181 (Cys in wild type enzyme). Kinetic data are given in Tables 2 and 3.

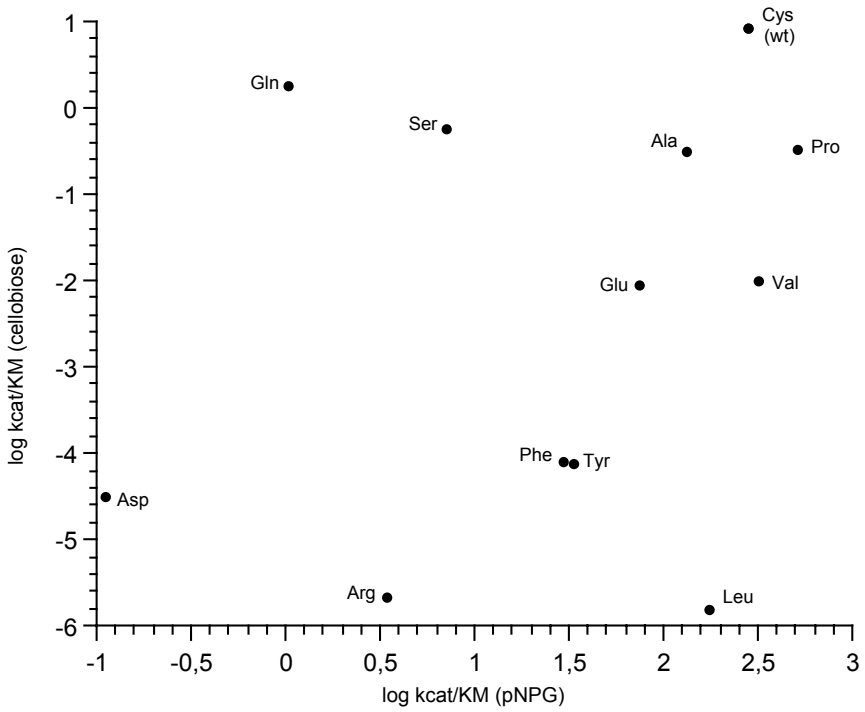



**Figure 5**. Efficiency and specificity. Log$k_{cat}/K_M$ relationship between cellobiose and p-nitrophenyl-β-D-glucoside substrates for Bgl3 mutants at position 181 (Cys in wild type enzyme). Kinetic data are given in Tables 2 and 3.