JOURNAL OF BACTERIOLOGY, Apr. 1995, p. 1981–1988 Vol. 177, No. 8
0021-9193/95/$04.00+0

Analysis of a Novel Gene and β-Galactosidase Isozyme from a Psychrotrophic *Arthrobacter* Isolate

KEVIN R. GUTSHALL, DONALD E. TRIMBUR,† JODIE J. KASMIR, AND JEAN E. BRENCHLEY*

Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania

Received 21 September 1994/Accepted 28 December 1994

We have characterized a new psychrotrophic *Arthrobacter* isolate which produces β-galactosidase isozymes. When DNA from this isolate was transformed into an *Escherichia coli* host, we obtained three different fragments, designated 12, 14, and 15, each encoding a different β-galactosidase isozyme. The β-galactosidase produced from fragment 12 was of special interest because the protein subunit was smaller (about 71 versus 116 kDa) than those typically encoded by the *lacZ* family. The isozyme encoded by fragment 12 was purified, and its activity and thermostability were examined. Although the enzyme is highly specific towards β-D-galactoside substrates, its levels in the isolate do not increase in cells grown with lactose. Nucleotide sequence determination showed that the gene encoding isozyme 12 is not similar to the other members of the *lacZ* family but has regions similar to β-galactosidase isozymes from *Bacillus stearothermophilus* and *B. circulans*. Addition of the isozyme 12 sequence to the database made it possible to examine these enzymes as possible members of a new, separate family. Our analysis of this new family showed some conserved amino acids corresponding to the *lacZ* acid-base catalytic region but no homology with the nucleophilic region. On the basis of these comparisons, we designated this a new *lacG* family.

The intricate genetic and biochemical studies with the *lacZ*-encoded β-galactosidase (EC 3.2.1.23) of *Escherichia coli* provide an excellent foundation for examining and comparing the structures of other β-galactosidases. Because the *lacZ* enzyme is used so frequently, we often overlook other interesting and unusual β-galactosidases. The study of these new enzymes, however, can help explain the roles of various amino acids in enzyme structure, provide insight into the evolution of genes encoding β-galactosidases, suggest structural relationships to other glycosidases, and yield enzymes with valuable properties that differ from those of the *E. coli* enzyme.

We had isolated psychrotrophic *Arthrobacter* strains which produced cold-active β-galactosidases and demonstrated that cells of one strain, B7, contained at least two isozymes when grown with lactose as a carbon source (22). When we used DNA from isolate B7 to transform an *E. coli* host, we obtained three transformants, each containing a different DNA segment encoding a β-galactosidase activity. One gene, designated 15, encoded the predominant isozyme found in lactose-grown cells. This enzyme has a temperature optimum about 20°C below that of the *E. coli* β-galactosidase and has been purified and characterized, and the gene that encodes it has been sequenced (33). This enzyme has the conserved acid-base and nucleophilic sites involved in catalysis that are typically found with the *lacZ* family of β-galactosidases. The presence of the other isozymes in this isolate raised questions about their possible functions in enabling these psychrotrophic strains to survive rapid temperature changes or in using other glycoside substrates. Our initial work with the second isozyme, designated 12, suggested that it was unlike other β-galactosidases reported and that its characterization would add insight into the biochemical properties and function of β-galactosidases.

* Corresponding author. Phone: (814) 863-7794. Fax: (814) 863-7024. Electronic mail address: jeb7@psu.edu.

† Present address: Genencor International, South San Francisco, Calif.

In this report, we describe the purification and characterization of this second isozyme, examination of its regulation in isolate B7, determination of the nucleotide sequence of the gene, and its comparison with other proteins reported in GenBank. This enzyme, designated isozyme 12, differs significantly from isozyme 15. Isozyme 12 has a temperature optimum between 45 and 50°C, which is closer to that found for the *E. coli* β-galactosidase. However, its subunit is only about 71 kDa, which is considerably smaller than that found for gene 15 (111 kDa) and other typical *lacZ* enzymes (116 kDa for *E. coli*). A comparison with other sequences in GenBank showed that the β-galactosidase from gene 12 shares homology with one isozyme reported from the thermophile *Bacillus stearothermophilus* and two isozymes found in the mesophile *B. circulans*. Our results obtained with isozyme 12 show that these enzymes constitute a new class of non-*lacZ* β-galactosidases. These enzymes share homology with the acid-base catalytic region of the *lacZ* family but do not have significant homology with the nucleophilic site. A comparison of the isozyme 12 sequence with the domains recently reported for the crystal structure of the *E. coli lacZ* β-galactosidase (18) show that it is not just a truncated version of the *lacZ* protein. The smaller size of these subunits, coupled with the difference in their nucleophilic sites, raises interesting questions about which structural features are required for β-galactosidase activity and how these novel enzymes might have evolved. Our results obtained with isozyme 12 also make it possible to define a new β-galactosidase family, which we have designated *lacG* to indicate the enzymes' galactosidase or possible glycosidase activity.

MATERIALS AND METHODS

Bacterial strains and media. Psychrotrophic *Arthrobacter* strain B7 was isolated from Pennsylvania farmlands as previously described (22). The strain was grown and maintained on tryptic soy agar at 25°C. *E. coli* JM109 (34) and MC1061 (23) were maintained on M9 minimal medium as described by Miller (24). *E. coli* strains containing plasmids pΔα18 and pUC18 (34) were grown in Luria broth (24) containing ampicillin (100 μg/ml). *B. stearothermophilus* IAM11001 was obtained from the American Type Culture Collection (Bethesda,

Md.) as ATCC 8005. It has been suggested that this strain be reclassified as *B. kaustophilus* (1, 26). Since the publications describing the enzymes and genes of interest refer to this as *B. stearothermophilus*, we will use this name here to avoid confusion. The *B. stearothermophilus* strain was grown on nutrient agar and broth (Difco) at 55°C.

DNA isolation. Both *Arthrobacter* isolate B7 and *B. stearothermophilus* DNAs were isolated by a modification of the method of Brahamsha and Greenberg (3). The lysozyme concentration was increased to 5 mg/ml with 1% (wt/vol) sodium dodecyl sulfate (SDS) and incubated at 60°C for 1 h.

Plasmid preparation. To screen for genes encoding β-galactosidase activity, we had to construct a derivative of pUC18 lacking the alpha portion of the *E. coli lacZ* gene. The *Nde*I-*Hin*dIII fragment was deleted from pUC18 to construct plasmid pΔα18 (33). Plasmid pUC18 was used for all subcloning of the DNA fragment encoding enzyme 12. Plasmid DNA was prepared by the method of Birnboim and Doly (1a).

Cloning protocols. Cloning of the DNA encoding isozyme 12 was accomplished by partially digesting isolate B7 DNA with *Sau*3AI, separating the fragments on a 0.4% agarose gel, excising fragments in the size range of 4 to 6 kb by using the USBioclean MP kit (United States Biochemical, Cleveland, Ohio), and ligating the fragments into plasmid pΔα18 DNA that had been digested with *Bam*HI. Competent cells of strain JM109 were transformed, and transformants capable of hydrolyzing X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) were selected as previously described (33). Transformants containing a DNA segment encoding isozyme 12 turned blue at 37°C, whereas those carrying gene 15 were white at the higher temperatures and turned blue when cells were incubated at 25°C or below. The original 3.7-kb fragment was subcloned to form a 2.2-kb segment designated clone 12.1 by treatment with *Eco*RI and *Pst*I. Clone 12.1 was used for purification of the isozyme 12 β-galactosidase activity and to derive the *Bam*HI-to-*Pst*I fragments used for Southern blot analysis and nucleotide sequencing. The region identified as encoding isozyme 12 by sequence analysis was designated gene 12.

Subcloning and sequencing. Subclones were prepared from clone 12.1 by digestion with restriction endonucleases, ligation into pUC18, and transformation of *E. coli* JM109 competent cells. DNA sequencing was performed by a modification of the method of Tabor and Richardson (31), with the Sequenase version 2.0 kit (United States Biochemical). Because *Arthrobacter* strains have a high G+C content, deaza-7-dGTP replaced dGTP and the reaction temperature was increased to 42°C to decrease compressions. For the subclones generated by restriction digests, the M13 reverse and universal primers were used and additional oligonucleotides were synthesized with a Millipore DNA Synthesizer 7500 (Millipore, Bedford, Mass.).

Southern analysis. Either plasmid DNA or DNA prepared from isolate B7 was digested with *Bam*HI for 2.5 h and separated on a 0.5% (wt/vol) agarose gel. DNA was transferred to an Immobilon-S membrane (Millipore) with a Trans-Blot apparatus (Bio-Rad, Melville, N.Y.). DNA was hybridized at 68°C for 16 h (30) with a biotin-labeled *Bam*HI-*Pst*I fragment obtained from subclone 12.1. The label was detected with the Polar Plex chemiluminescent blotting kit (Millipore).

Enzyme purification. A transformant of *E. coli* MC1061 containing plasmid pΔα18 with DNA fragment 12.1 inserted was grown in 500 ml of TYP medium (containing, per liter, 16 g of tryptone, 16 g of yeast extract, 5 g of NaCl, and 2.5 g of K_2HPO_4) supplemented with 100 μg of ampicillin per ml. The culture was incubated at 25°C for 36 h. Cells were harvested by centrifugation at 9,800 × *g* for 10 min at 4°C. The cells were washed and resuspended in Z buffer (24) and then passed through a French pressure cell at 14,000 lb/in^2.

DNase (100 μg) was added to the cellular extract. After incubation on ice for 10 min, the extract was centrifuged at 30,000 × *g* for 20 min at 4°C. The supernatant was collected and precipitated with 20% (wt/vol) ammonium sulfate at 0°C for 30 min and then centrifuged at 30,000 × *g* for 30 min at 0°C. The supernatant was collected, ammonium sulfate was added to a final concentration of 40%, and the mixture was incubated at 0°C for 30 min. The extract was centrifuged as before, the supernatant was discarded, and the pellet, which contained the activity, was resuspended in 20 mM sodium phosphate buffer (pH 7.0). The extract was diluted to 120 ml with 20 mM $NaPO_4$ buffer (pH 7.0) and applied to a 20-ml bed volume of Macro Q anion-exchange resin (Bio-Rad Laboratories) at a flow rate of 1.0 ml/min at 4°C. Following sample loading, the column was washed with 2 bed volumes of 20 mM $NaPO_4$–200 mM NaCl buffer (pH 7.0). The enzyme was eluted with a 100-ml linear gradient of 200 to 400 mM NaCl–20 mM $NaPO_4$ buffer (pH 7.0). The flow rate was 0.1 ml/min, and 2.0-ml samples were collected. Samples containing activity were pooled and concentrated with the addition of 50% ammonium sulfate and placed on ice for 30 min. The sample was centrifuged at 30,000 × *g* for 15 min, and the supernatant was discarded. The pellet was resuspended in 1.0 ml of Z buffer and desalted on a 10-ml Sephadex G-25 column. The samples which contained activity were applied to a 20-ml affinity matrix of agarose derivatized with *p*-aminobenzyl-1-thio-β-D-galactopyranoside (Sigma, St. Louis, Mo.), washed with 20 mM $NaPO_4$–200 mM NaCl buffer (pH 7.0), and then eluted with 20 mM $NaPO_4$–400 mM NaCl buffer (pH 7.0). As before, the active samples were pooled, concentrated, and desalted. The N-terminal amino acid sequence was determined at the Hershey Medical Center of The Pennsylvania State University.

Enzyme assays. β-Galactosidase activity was measured as the release of *o*-nitrophenol from *o*-nitrophenyl-β-D-galactopyranoside (ONPG) at 420 nm on a Hewlett-Packard diode array spectrophotometer (24). One unit of activity is defined as 1 μmol of *o*-nitrophenyl released per min. Specific activity is defined as units per milligram of protein. Assays were performed by adding the enzyme to Z buffer which contained 0.5 M β-mercaptoethanol. The reaction was started by addition of the substrate, ONPG, and stopped by addition of 1 M Na_2CO_3. Assays using lactose as a substrate were started by addition of various concentrations of lactose. The reaction was stopped by boiling the sample in a water bath for 3 min. The galactose dehydrogenase assay was used to measure the amount of galactose released by the enzyme (27). The specific activity is defined as micromoles of galactose released per minute per milligram of protein. Protein concentrations were determined by the method of Bradford (2), with the Bio-Rad protein assay dye reagent concentrate. Standard curves were generated with bovine serum albumin.

Polyacrylamide gel electrophoresis and Western blot (immunoblot) analysis. Proteins were analyzed by the method of Laemmli (20) on 7.5% (wt/vol) polyacrylamide gels and stained with Coomassie blue. For Western blots, polyclonal antibodies were produced against the recombinant proteins purified from genes 15 and 12 by injecting the proteins emulsified with Freund's complete adjuvant into New Zealand White rabbits. Serum was collected after 6 weeks and absorbed with crude extracts of *E. coli* MC1061 cells. Western analysis was done by the method of Towbin et al. (32), with goat anti-rabbit immunoglobulin G conjugated to alkaline phosphatase and detected with BCIP (5-bromo-4-chloro-3-indolylphosphate) and nitroblue tetrazolium (21).

Chemicals and reagents. Components of the media were from Difco. Restriction endonucleases, ligases, polymerase, and nucleotides were purchased from Gibco-BRL (Gaithersburg, Md.) and Boehringer Mannheim (Indianapolis, Ind.). Western blot components, as well as ONPG and X-Gal, were purchased from Sigma (St. Louis, Mo.). Radionucleotides were purchased through New England Nuclear Corporation.

Nucleotide sequence accession number. The GenBank accession number for the gene 12 sequence from isolate B7 reported here is U17417.

RESULTS

Southern analysis of fragment 12. Because three DNA segments with different restriction endonuclease sites were found when DNA from isolate B7 was transformed into *E. coli*, it was important to establish that each of these originated from strain B7 and not from some other source. To confirm that DNA fragment 12 was in fact *Arthrobacter* B7 DNA, Southern blot hybridizations were performed. A *Bam*HI-*Pst*I fragment obtained from subclone 12.1 was used to probe against DNAs from the *E. coli* host strain, *Arthrobacter* isolate B7, and the three transformants carrying plasmids with cloned fragments 12, 14, and 15 (Fig. 1). The results showed that the probe hybridized with DNA from isolate B7 and the transformant carrying fragment 12. There was no hybridization with the *E. coli* host strain or fragments 14 and 15 in the other transformants.

Purification of the isozyme synthesized from gene 12. Polyacrylamide gel electrophoresis demonstrated that the protein made in transformants carrying fragment 12 migrated to the same position as the second isozyme detected in isolate B7 cells grown with lactose (33). Since this isozyme represents only a few percent of the β-galactosidase activity in isolate B7 (the remaining amount corresponds to the protein expressed by gene 15), it was difficult to purify this minor isozyme directly from isolate B7. The *E. coli* transformant carrying subclone 12.1, however, had a high concentration of the enzyme (32.3 U/mg). Thus, this enzyme, referred to as isozyme 12, was purified from the *E. coli* transformant by 20 to 40% ammonium sulfate precipitation, anion-exchange chromatography, and elution from an affinity column as described in Materials and Methods and shown in Table 1 and Fig. 2. The final preparation was at least 90% pure and was used for antibody generation and enzyme characterization. N-terminal amino acid analysis (see the later section on the nucleotide sequence of gene 12) showed that the first eight amino acids may have been cleaved from the full-length protein and that a slightly truncated enzyme was characterized.

Antibodies were used in Western blot analysis to establish that the enzyme from the *E. coli* transformant migrated at the

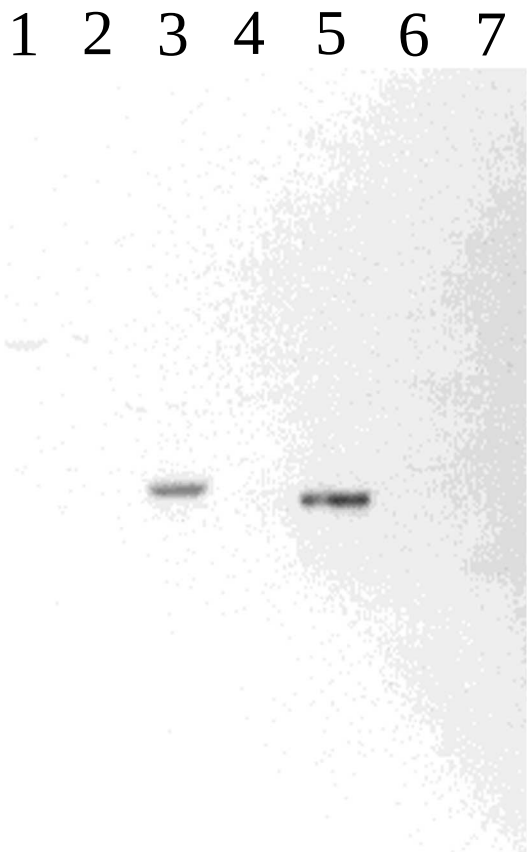

FIG. 1. Southern blot analysis with the DNA fragment encoding isozyme 12. Lanes 1, 2, and 3 contained pΔα plasmid DNAs carrying β-galactosidase-encoding fragments 15, 14, and 12, respectively. These plasmids were prepared from the *E. coli* transformants and digested with *Bam*HI. Lanes 4 through 7 contained *Bam*HI-digested DNAs prepared from *E. coli* JM109 and psychrotrophic isolates B7, D5, and D2, respectively. The probe was an internal *Bam*HI-*Pst*I fragment of subclone 12.1. Lanes: 1, pΔαB7-15; 2, pΔαB7-14; 3, pΔαB7-12; 4, *E. coli* JM109; 5, isolate B7; 6, isolate D2; 7, isolate D5.

same position during polyacrylamide gel electrophoresis as another cross-reacting protein in the original isolate. These results are consistent with the notion that the enzyme synthesized by the *E. coli* transformant is the same as the isozyme found in *Arthrobacter* isolate B7 (Fig. 2). The apparent subunit molecular mass of the purified isozyme, on the basis of SDS-polyacrylamide gel electrophoresis, was about 71 kDa (Fig. 2).

Characterization of isozyme 12. To determine the optimal pH for subsequent studies, enzyme activity was measured in 50 mM Tris buffers with pH values of 7.7 to 9.0 and in 100 mM sodium phosphate buffers with pH values of 6.0 to 8.0. The enzyme maintained activity over a broad range of pH values, from 6 to 9, but showed a peak at pH 6.6 (data not shown). The effects of sodium and potassium salts were determined by assaying activity in 10 mM phosphate buffer containing sodium and potassium at concentrations of 0 to 450 mM. Optimal activity occurred at concentrations between 100 and 250 mM, and no difference between sodium and potassium was observed (data not shown). For determination of other ion effects, the enzyme was first treated with 20 mM EDTA and desalted by G-25 Sephadex column chromatography. The enzyme was eluted from the column with 100 mM morpholinepropanesulfonic acid (MOPS) buffer (pH 6.6). Even after 2 h of EDTA treatment, about 26% of the original activity remained once the EDTA was removed. Addition of Mg^{2+} at 5 mM and Mn^{2+} or Ca^{2+} at 10 mM restored activity (data not shown). Even addition of Co^{2+} increased activity twofold, suggesting that the ion effect was not highly specific; however, 10 mM Ni^{2+} had no effect and 10 mM Cu^{2+} inhibited activity (data not shown). Addition of β-mercaptoethanol also stimulated activity. Therefore, the buffer selected for subsequent characterization of enzyme activity consisted of 100 mM KPO_4 (pH 6.6), 100 mM KCl, 5 mM MgCl, and 5 mM β-mercaptoethanol.

The effects of temperature on isozyme 12 were examined in two ways. The first was to determine the temperature optimum for catalysis by assaying the enzyme at temperatures of 0 to 60°C. The results showed a linear increase in activity with increasing temperature, with the maximum activity between 45 and 50°C and a rapid decline above 50°C (data not shown). In the second analysis, the thermostability of the activity was determined by incubating the enzyme in buffer at various temperatures, removing samples at different times, and assaying enzyme activity at 30°C with ONPG as the substrate (Fig. 3). The results showed that the enzyme was quite stable for at least 70 h at temperatures of 35°C and below. However, at 50°C the enzyme lost all activity in less than 15 min. Even though isozyme 12 had a higher optimum temperature than isozyme 15, both were readily inactivated at 50°C.

Assays were performed with ONPG or lactose at various concentrations as a substrate to determine the K_m and V_{max} values. Isozyme 12 had an apparent K_m of 0.57 mM and a V_{max} value of 254 U/mg of protein with ONPG at concentrations of 0.05 to 5.0 mM. For lactose, the apparent K_m was 4.81 mM and the V_{max} value was 3.97 U/mg of protein when assayed at concentrations of 0.1 to 50 mM. The *E. coli lacZ* β-galactosidase has apparent K_m values of 1.4 and 0.1 and V_{max} values of 31 and 390 U/mg of protein with lactose and ONPG, respectively.

We considered the possibility that isozyme 12 might hydrolyze substrates other than ONPG or lactose. To examine its substrate specificity, the enzyme was assayed by using 11 *p*-nitrophenyl (NP)-glycoside substrates at a final concentration of 0.5 mM at 30°C (Table 2). The only significant activity was seen when *p*NP-β-D-galactoside was used. Thus, if this isozyme functions on other substrates, this activity was not apparent with any of the typical glycoside compounds.

Regulation of isozymes 12 and 15 in isolate B7. The availability of antibodies to purified isozymes 12 and 15 made it possible to use Western blot analysis to examine the regulation

TABLE 1. Purification of *Arthrobacter* isolate B7 β-galactosidase expressed by clone 12.1

Step	Vol (ml)	Total protein (mg)	Total activity (μmol/min)	Sp act (μmol/min/mg)	Recovery (%)	Purification (fold)
Crude extract	22.0	473.0	15,290.0	32.3	100.0	1.0
Ammonium sulfate (20–40%)	25.0	180.0	14,725.0	81.8	96.3	2.5
Anion-exchange chromatography	17.1	59.5	12,483.0	209.8	81.6	6.5
Affinity chromatography	24.0	28.8	8,880.0	308.3	58.1	9.5

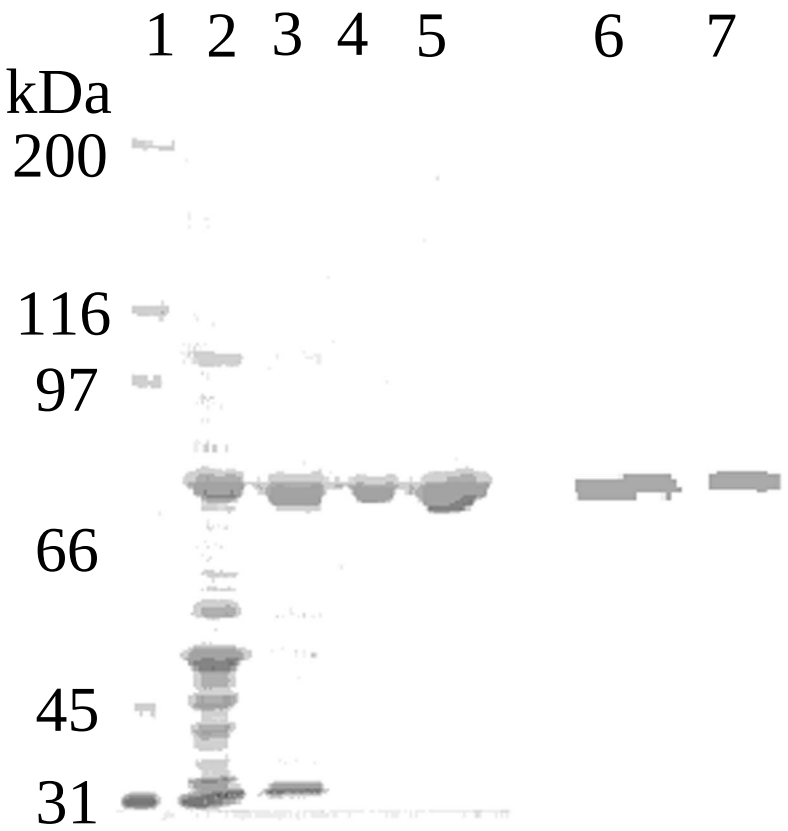

FIG. 2. Purification and Western blot analysis of *Arthrobacter* isolate B7 β-galactosidase isozyme 12. Isozyme 12 was purified from an *E. coli* transformant containing pΔαB7-12.1, and samples from each step were subjected to electrophoresis on an SDS–7.5% polyacrylamide gel and stained with Coomassie blue. Lanes: 1, molecular weight markers; 2, crude extract (2.0 μg of protein); 3, ammonium sulfate precipitation (0.79 μg of protein); 4, anion-exchange chromatography (0.31 μg of protein); 5, affinity chromatography (0.21 μg of protein); 6, Western blot analysis with antibody to isozyme 12 reacted against crude extract from isolate B7; 7, Western blot analysis with antibody to isozyme 12 reacted against crude extract from the *E. coli* MC1061 transformant containing plasmid pΔαB7-12.1.

of these isozymes in the original isolate, B7. Since isolate B7 grows on cellobiose, we considered the possibility that isozyme 12 might be regulated by this carbon source even though the in vitro enzyme assays suggested that cellobiose was not a good substrate. Since the isozymes had slightly different temperature profiles, we grew cells at a low temperature (4°C) and a temperature near the maximum for growth (25°C) to determine whether the expression of either isozyme might be regulated differently at these growth temperatures. Isolate B7 cells were grown at 4 and 25°C in M9 medium with either lactose, cellobiose, or glucose at 0.2% as a carbon source. Cells were assayed for total β-galactosidase activity (Table 3), and extracts were prepared and used for Western blot analysis (Fig. 4). The total β-galactosidase activity was increased 20- to 60-fold in

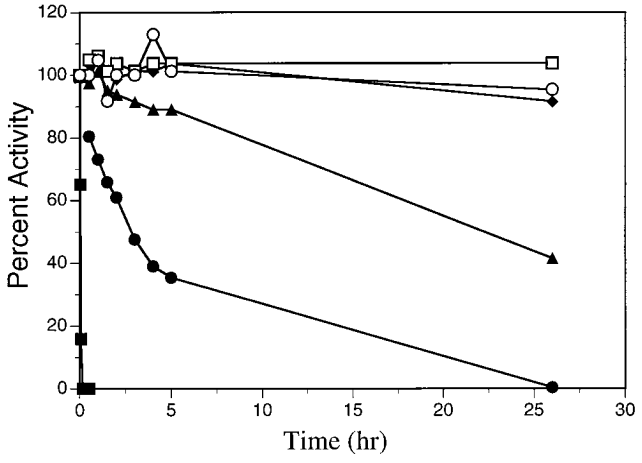

FIG. 3. Thermostability of β-galactosidase isozyme 12 from *Arthrobacter* isolate B7. The enzyme was incubated at different temperatures, and samples were removed at various times and assayed for activity at 30°C with ONPG as the substrate. Symbols: ■, 50°C; ●, 45°C; ▲, 40°C; ◆, 35°C; □, 30°C; ○, 25°C.

TABLE 2. Substrate specificity of isozyme 12 from isolate B7

Substrate	% Activity	
	Crude extract	Purified isozyme 12
pNP-β-D-galactoside	100.0[a]	100.0[a]
pNP-α-D-galactoside	0.2	<0.01
pNP-β-D-glucoside	0.01	<0.01
pNP-β-D-mannoside	0.2	<0.01
pNP-β-L-fucoside	0.7	0.45
pNP-β-D-xyloside	<0.01	<0.01
pNP-β-L-arabinoside	0.1	<0.01
pNP-β-D-galuronide	0.3	<0.01
pNP-β-D-glucuronide	0.1	<0.01
pNP-β-D-lactoside	0.2	<0.01
pNP-β-D-cellobioside	0.2	<0.01

[a] The 100% values are 1.1 and 130 U/mg for the crude extract and purified isozyme 12, respectively.

TABLE 3. Regulation of β-galactosidase activity in isolate B7 cells grown with different carbon sources at 25 and 4°C[a]

Carbon source	Growth temperature (°C)	Sp act (μmol/min/mg of protein)
Lactose	25	1.84
Cellobiose	25	0.05
Glucose	25	0.03
Lactose	4	0.92
Cellobiose	4	0.05
Glucose	4	0.09

[a] Cells were grown at the specified temperatures in M9 medium containing the indicated carbon source at 0.2%. Specific activity was measured with ONPG as the substrate.

cells grown with lactose as opposed to glucose and cellobiose (Table 3). The Western blots showed that isozyme 15 increased dramatically in lactose-grown cells. Isozyme 15 accounted for the increased β-galactosidase activity, whereas isozyme 12 levels remained relatively constant for cells grown in all three media (Fig. 4).

Nucleotide sequence of gene 12. The original 3.7-kb fragment containing gene 12 was further subcloned to a 2.2-kb (*Eco*RI-*Pst*I) segment, subclone 12.1, which still produced the active enzyme. A restriction endonuclease map was generated from subclone 12.1 with enzymes recognizing six base pairs, and further subclones were prepared for sequencing. Because of the lack of frequent restriction endonuclease sites, oligonucleotide primers were synthesized and used to complete the double-stranded sequencing from the *Bam*HI to the *Pst*I sites (Fig. 5). The start site (ATG) is at nucleotide position 119 and is followed by an open reading frame of 1,914 nucleotides. The stop codon (TGA) ends at nucleotide position 2033. Comparison of the upstream region with that found for gene 15 revealed no obvious regulatory regions other than a possible ribosome binding site (Fig. 5).

The reading frame was confirmed by determination of the N-terminal sequence of the recombinant B7 12.1 protein. The isozyme 12 protein was immobilized on a polyvinylidene difluoride membrane and sequenced by Edman degradation on an ABI 477A/120A analyzer. The first eight cycles revealed an N-terminal sequence of E-G-T-T-P-D-T-A. This sequence is eight amino acids beyond the methionine start site predicted from the nucleotide sequence. We examined the sequence for all three possible start codons, ATG, GTG, and TTG. The only open reading frame capable of encoding a 71-kDa protein begins with the ATG at nucleotide 119. Since the results of the

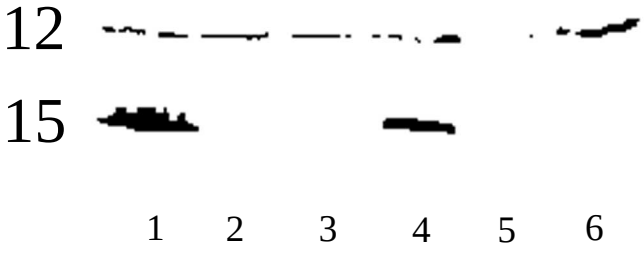

FIG. 4. Western blot analysis of the expression of isozymes 12 and 15 in isolate B7 grown with different carbon sources. All lanes contained 100 μg of protein from isolate B7 crude extracts. Cells were grown with the different carbon sources at either 25 or 4°C. The blots were reacted with antibodies to isozymes 12 (top row) and 15 (bottom row). Lanes: 1, lactose at 25°C; 2, cellobiose at 25°C; 3, glucose at 25°C; 4, lactose at 4°C; 5, cellobiose at 4°C; 6, glucose at 4°C.

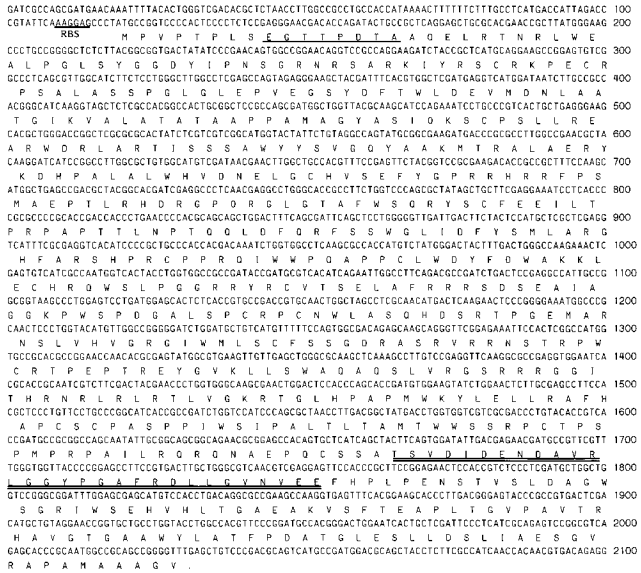

FIG. 5. The nucleotide sequence of the 2.2-kb *Bam*HI-*Pst*I fragment containing the gene encoding β-galactosidase isozyme 12. The deduced amino acid sequence for isozyme 12 is included below the nucleotide sequence. The putative ribosome binding site (RBS) and the amino acids found during determination of the N-terminal sequence of the purified protein are underlined. The double underlining highlights amino acids 516 to 545. This region is highly conserved in isozyme 12 and the three *Bacillus* isozymes and may contain a nucleophilic site.

N-terminal amino acid sequencing establish the reading frame, it is likely that translation begins with the methionine noted and that a short peptide fragment is cleaved in the *E. coli* transformant by protease activity.

Comparison of the amino acid sequence of isozyme 12 with those of other β-galactosidases. The translated amino acid sequence from gene 12 was compared to that found for gene 15 and others available in GenBank. There was little similarity to isozyme 15 or other members of the *lacZ* family of β-galactosidases. Interestingly, the sequence which showed the most similarity (Table 4) was one of the β-galactosidase isozymes obtained from the thermophile *B. stearothermophilus* (15–17). To determine whether we could confirm this similarity at the protein level, we used an antibody prepared against isozyme 12 in a Western blot analysis to see if it cross-reacted to any protein in *B. stearothermophilus*. The results (Fig. 6) show that the antibody cross-reacted with a *B. stearothermophilus* protein. There was no cross-reaction with extracts from *E. coli* transformants carrying either of the other two isolate B7 genes, 15 and 14.

We also examined the alignment of the conserved regions reported to be involved in the acid-base and nucleophilic transfer during catalysis by β-galactosidases. The proposed mechanism of action for the *E. coli lacZ* β-galactosidase involves a double-displacement reaction in which the enzyme forms and hydrolyzes a glycosyl-enzyme intermediate via oxocarbonium ion-like transition states (8). In their studies on the mechanism of action of the *E. coli lacZ* β-galactosidase, Gebler et al. (8) presented evidence that the glutamate at residue 537 is the nucleophilic site. Alignment of other *lacZ* β-galactosidases showed significant homology surrounding this glutamate, forming the consensus sequence I-L-C-E-Y-A-H-A-M-G-N. A second glutamate at residue 461 was proposed as the general acid-base catalyst which is important to the binding of Mg^{2+}

and which protonates the leaving group and deprotonates the attacking water (8). This view is supported by the highly conserved region surrounding this glutamate when the sequences for *lacZ* β-galactosidases are aligned (33).

The comparison of the sequence of isozyme 12 failed to find any homology with the consensus nucleophilic region of the *lacZ* family. However, a possible alignment was found for the proposed acid-base site (Fig. 7). Although isozyme 12 and the *Bacillus* enzymes showed significant homology with those in the *lacZ* family at key amino acids in the acid-base region, the overall conservation is not as great. These isozymes may have evolved their similar functions separately from the *lacZ* β-galactosidases.

DISCUSSION

In our search for cold-active enzymes, we have discovered a unique *Arthrobacter* strain that produces two β-galactosidases in lactose-grown cells and carries at least three β-galactosidase-encoding genes. Our Southern hybridization results established that all three genes originated from *Arthrobacter* isolate B7. The finding of these three genes illustrates the power of using recombinant DNA technology to discover genes that may be silent, are expressed at low levels, or make enzymes that are masked by similar activities during the screening process. Because these genes may encode proteins with unusual and interesting properties, we are examining each one to explore its in vivo function and potential industrial use.

Another intriguing feature of this work is the prevalence of β-galactosidase isozymes found in various microorganisms. For example, *B. stearothermophilus* has two genes producing β-galactosidase isozymes, and the sequences of two isozymes from *B. circulans* are recorded in GenBank. It could be that isozymes are more common than anticipated, especially in extremophiles, where they may help to adapt to rapidly changing growth conditions. It is also possible that the highly sensitive chromogenic substrates used in these studies readily detect enzymes with related activities and unknown functions. In addition, we considered the possibility that β-galactosidase genes may be common because they are plasmid encoded and readily transferred to different organisms. For example, Guiso and Ullmann (10) studied several plasmids which transferred genes for lactose degradation among members of the family *Enterobacteriaceae* and plasmid-encoded β-galactosidases have been reported for a *Klebsiella* strain (4, 11). We considered this possibility and did preliminary experiments to detect plasmids in isolate B7. Although our results would not rule out the existence of large or very low copy number plasmids, there is no evidence to suggest that these genes are plasmid borne.

Our previous analysis of gene 15 had shown that it has the conserved active sites associated with the *lacZ* family of β-galactosidases (33). Here we used Western blot analysis (Fig. 4) to show that isozyme 15 is regulated as a typical lactose utilization enzyme. Its levels increase in cells grown with lactose and decrease in cells grown with either glucose or cellobiose. These results suggest that isozyme 15 is responsible for lactose utilization under the growth conditions studied. Isozyme 12, however, is present at low levels and does not show significant regulation in cells grown with the three carbon sources or at 4 or 25°C (Fig. 4). It is not clear whether isozyme 12 has a direct role in lactose utilization in isolate B7, but *E. coli* transformants expressing isozyme 12 form colonies on media containing lactose as a carbon source. In addition, the enzyme is highly specific for β-D-galactosides and no other substrate has been identified (Table 2). It is possible that isozyme 12 functions

TABLE 4. Amino acid sequence similarity of β-galactosidases[a]

Enzyme source	% Similarity to enzyme from:													
	Arthrobacter isolate B7 12	*B. stearothermophilus*	*B. circulans bgaA*	*B. circulans bgaB*	*S. solfataricus* p22498	*S. solfataricus* p14288	*Arthrobacter* isolate B7 15	*E. coli lacZ*	*E. coli ebgA*	*K. pneumoniae*	*C. acetobutylicum*	*S. thermophilus*	*L. bulgaricus*	*L. lactis*	*K. lactis*
Arthrobacter isolate B7 12	100.0	21.8	17.7	17.6	10.8	11.2	11.6	11.6	9.4	10.7	9.2	9.9	10.2	10.8	9.2
B. stearothermophilus		100.0	41.5	36.3	11.0	10.0	10.6	10.1	9.8	10.1	10.0	10.0	11.0	10.3	10.7
B. circulans bgaA			100.0	41.3	10.2	10.4	11.1	11.7	10.4	10.5	10.5	11.1	12.7	10.5	10.5
B. circulans bgaB				100.0	11.2	12.0	11.3	10.4	11.4	10.9	10.0	11.0	11.9	10.1	10.7
S. solfataricus p22498					100.0	71.4	11.0	10.4	10.2	11.5	11.5	11.9	12.1	10.8	10.2
S. solfataricus p14288						100.0	11.6	11.2	11.1	12.0	10.4	10.4	11.0	11.6	10.8
Arthrobacter isolate B7 15							100.0	25.7	26.0	26.0	21.9	25.6	20.0	18.7	24.4
E. coli lacZ								100.0	27.6	57.7	28.5	27.2	26.9	24.4	23.9
E. coli ebgA									100.0	26.1	24.4	21.7	23.3	19.7	28.3
K. pneumoniae										100.0	23.9	26.6	22.2	22.5	23.8
C. acetobutylicum											100.0	47.7	39.0	36.3	22.1
S. thermophilus												100.0	45.3	35.1	22.8
L. bulgaricus													100.0	30.4	19.4
L. lactis														100.0	16.9
K. lactis															100.0

[a] Sequence comparison was performed by using the Clustal method with a K-tuple of 1 and a gap penalty of 10. Nonidentical residue substitutions were weighted according to PAM 250 settings.

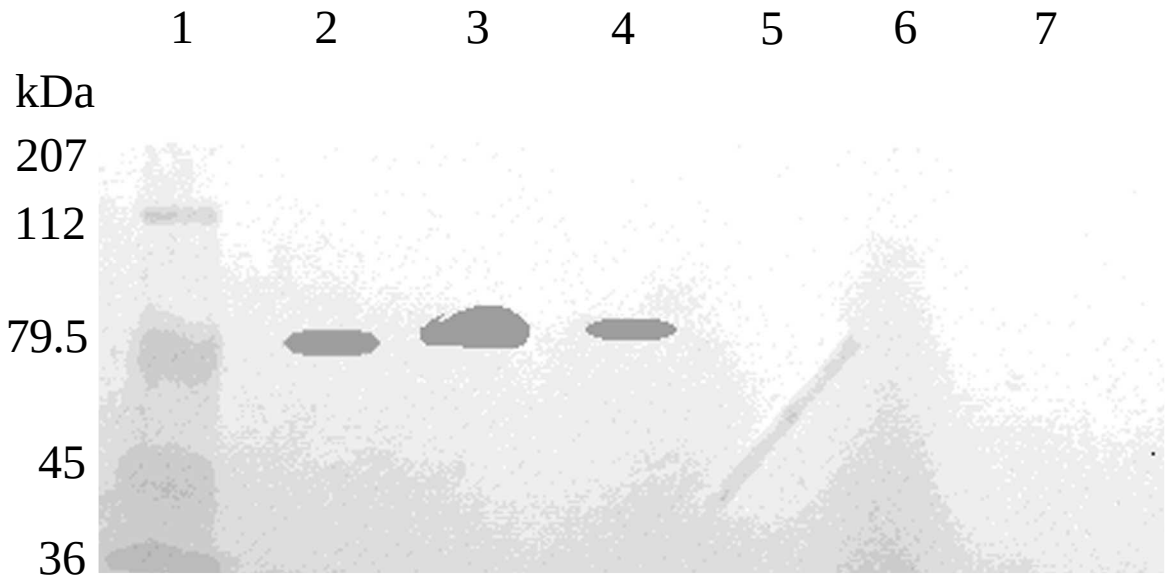

FIG. 6. Western blot analysis of antibody prepared against isozyme 12 reacted with crude extracts from *Arthrobacter* isolate B7, *E. coli* transformants, and *B. stearothermophilus*. The numbers on the left indicate the migration of molecular weight markers. Lanes: 1, molecular weight markers; 2, isolate B7; 3, *E. coli* JM109 containing pUC18 with clone 12.1 encoding isozyme 12; 4, *B. stearothermophilus*; 5, *E. coli* JM109 containing pUC18 with clone 15.12 encoding isozyme 15; 6, *E. coli* JM109 containing pUC18 with clone 14.2 encoding isozyme 14; 7, *E. coli* JM109.

during rapid temperature changes, during growth conditions not yet identified, or on some untested substrate.

The subunit sequence for isozyme 12 is unique in that it possesses the acid-base active site typical of the *lacZ* β-galactosidase family, but it does not share the conserved nucleophilic site. The smaller size of the subunit of isozyme 12 raised the question of whether it contains some, but not all, of the portions of the typical *lacZ* protein or whether it represents a novel structure with an analogous function. The recent publication of the three-dimensional structure of the *E. coli lacZ* β-galactosidase (18) made it possible to use the Clustal method to align the sequence of the subunit for isozyme 12 with each of the five domains found in the *E. coli* enzyme. Other than the homology found in domain 3 at the acid-base region, no areas of significant homology were found. This strongly suggests that isozyme 12 is not simply a truncated or reshuffled version of the *lacZ* domains.

Since isozyme 12 does not fit the *lacZ* family, we examined other glycosidases to determine whether it resembles any other

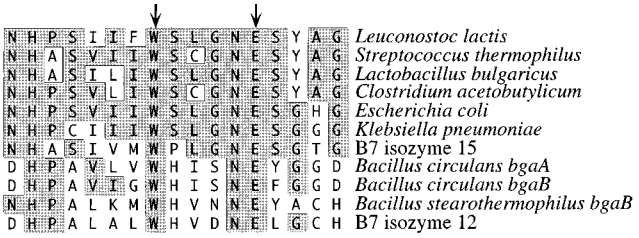

FIG. 7. Alignment of β-galactosidase acid-base active sites. Amino acid sequence alignment of β-galactosidases from *Arthrobacter* isolate B7 isozyme 15 (33), *E. coli lacZ* (19), *Klebsiella pneumoniae lacZ* (5), *Kluyveromyces lactis* (25), *Clostridium acetobutylicum* (12), *Streptococcus thermophilus* (29), *Lactobacillus bulgaricus* (28), *Leuconostoc lactis lacL* and *lacM* (7), *B. circulans bgaA* (GenBank accession number L03424), *B. circulans bgaB* (GenBank accession number L03425), *B. stearothermophilus* (15), and *Arthrobacter* B7 isozyme 12 (this report). Consensus residues are boxed. Proposed active-site residues are indicated by arrows.

known family. In his classification of glycosyl hydrolases, Henrissat (13) designated a separate class, *lacS* (6), for a β-galactosidase and a possible general glycosidase (9) from *Sulfolobus solfataricus*. Comparison of the sequence for isozyme 12 with those reported for *S. solfataricus* showed only 10 to 11% similarity to isozyme 12, showing that it is not in the *lacS* family (Table 4).

Our search of sequences in GenBank, however, yielded other intriguing similarities, especially to the smaller but more thermostable β-galactosidase isozyme found in the thermophile *B. stearothermophilus* (Table 4). Because of the homology suggested by the computer search, we examined the ability of an antibody prepared against isozyme 12 to cross-react with a protein(s) in *B. stearothermophilus*. Western blot analysis demonstrated that, indeed, a protein with a molecular mass corresponding to that reported for the *bgaB* gene (78 kDa) reacted with the antibody to isozyme 12. This antibody did not react with extracts from strains carrying the other *Arthrobacter* β-galactosidase genes, 14 and 15.

Our sequence comparisons also suggested some similarity to two isozymes from *B. circulans* (Table 4). Although their sequences are available in GenBank, no information on these enzymes has been published. The addition of our isozyme 12 sequence provided the new data needed to compare these sequences and determine whether they might form a separate family. The analysis showed some conservation of the acid-base site (Fig. 7). One other highly conserved region, corresponding to amino acids between positions 516 and 545 in isozyme 12 (Fig. 5), contains glutamate residues that might serve as a nucleophilic site. As new β-galactosidases are characterized, it will be useful to search for this proposed nucleophilic region to determine if it is highly conserved in other proteins. Henrissat (14) has shown that such searches not only highlight important amino acids and help identify active sites but can also suggest sequencing inaccuracies.

These results suggest that a new class of enzymes with β-galactosidase activity exists. We designated this group *lacG* to separate it from the *lacZ* and *lacS* families. If the *lacG* enzymes

have the same catalytic mechanism as those in the *lacZ* family, then it will be interesting to determine how the function of the nucleophilic site is accomplished. It is also possible that the catalytic mechanism of the *lacG* family differs from that found for the *lacZ* members. The discovery of proteins only 60 to 70% of the size of the *E. coli lacZ* β-galactosidase raises the question of why the *lacZ* subunits remained so large if a smaller protein can perform the same function. Comparisons of the *lacG* structures with that of the *E. coli lacZ* β-galactosidase should help clarify the role of the additional amino acids and determine which ones are important to the surface areas surrounding the substrate binding pocket and metal binding site. Since the *lacZ* and *lacG* families differ at the nucleophilic site, the possible similarity found at the acid-base region may reflect convergent rather than divergent evolution of these functions. It is likely that additional members of the *lacG* family exist, and their discovery and analysis should provide insight into the evolution and catalytic mechanism of these β-galactosidase activities.

ACKNOWLEDGMENT

We thank G. Farber and T. Nixon for helpful discussions. We thank our laboratory coworkers for suggestions, especially J. Loveland, who helped with data analysis.

This work was supported by Department of Energy grant DE-FG02-93ER20117 from the Division of Energy Biosciences and by National Science Foundation grant MCB-9306764.

REFERENCES

1. **Anonymous.** 1989. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. Int. J. Syst. Bacteriol. **39:**93–94.
1a. **Birnboim, H. L., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening of recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
2. **Bradford, 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.
3. **Brahamsha, B., and E. P. Greenberg.** 1987. Complementation of a *trpE* deletion in *Escherichia coli* by *Spirochaeta aurantia* DNA encoding anthranilate synthetase component I activity. J. Bacteriol. **169:**3764–3769.
4. **Brenchley, J. E., and B. Magasanik.** 1972. *Klebsiella aerogenes* strain carrying drug-resistance determinants and a *lac* plasmid. J. Bacteriol. **112:**200–205.
5. **Buvinger, W. E., and M. Riley.** 1985. Nucleotide sequence of *Klebsiella pneumoniae lac* genes. J. Bacteriol. **163:**850–857.
6. **Cubellis, M. V., C. Rozzo, P. Montecucchi, and M. Rossi.** 1990. Isolation and sequencing of a new β-galactosidase-encoding archaebacterial gene. Gene **94:**89–94.
7. **David, S., H. Stevens, M. van Riel, G. Simons, and W. M. de Vos.** 1992. *Leuconstoc lactis* β-galactosidase is encoded by two overlapping genes. J. Bacteriol. **174:**4475–4481.
8. **Gebler, J. C., R. Aebersold, and S. G. Withers.** 1992. Glu-537, not Glu-461, is the nucleophile in the active site of (lac Z) beta-galactosidase from *Escherichia coli*. J. Biol. Chem. **267:**11126–11130.
9. **Grogan, D. W.** 1991. Evidence that β-galactosidase of *Sulfolobus solfataricus* in only one of several activities of a thermostable β-D-glycosidase. Appl. Environ. Microbiol. **57:**1644–1649.
10. **Guiso, N., and A. Ullmann.** 1976. Expression and regulation of lactose genes carried by plasmids. J. Bacteriol. **127:**691–697.
11. **Hall, B. G., and E. C. R. Reeve.** 1977. A third β-galactosidase in a strain of *Klebsiella* that possesses two *lac* genes. J. Bacteriol. **132:**219–223.
12. **Hancock, K. R., E. Rockman, C. A. Young, L. Pearce, I. S. Maddox, and D. B. Scott.** 1991. Expression and nucleotide sequence of the *Clostridium acetobutylicum* β-galactosidase gene cloned in *Escherichia coli*. J. Bacteriol. **173:** 3084–3095.
13. **Henrissat, B.** 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280:**309–316.
14. **Henrissat, B.** 1993. Hidden domains and active site residues in β-glycanase-encoding gene sequences? Gene **125:**199–204.
15. **Hirata, H., T. Fukazawa, S. Negoro, and H. Okada.** 1986. Structure of a β-galactosidase gene of *Bacillus stearothermophilus*. J. Bacteriol. **166:**722–727.
16. **Hirata, H., S. Negoro, and H. Okada.** 1984. Molecular basis of isozyme formation of β-galactosidases in *Bacillus stearothermophilus*: isolation of two β-galactosidase genes, *bgaA* and *bgaB*. J. Bacteriol. **160:**9–14.
17. **Hirata, H., S. Negoro, and H. Okada.** 1985. High production of thermostable β-galactosidase of *Bacillus stearothermophilus* in *Bacillus subtilis*. Appl. Environ. Microbiol. **49:**1547–1549.
18. **Jacobson, R. H., X.-J. Zhang, R. F. DuBose, and B. W. Matthews.** 1994. Three-dimensional structure of β-galactosidase from *E. coli*. Nature (London) **369:**761–766.
19. **Kalnins, A., K. Otto, U. Ruther, and B. Muller-Hall.** 1983. Sequence of the *lacZ* gene of *Escherichia coli*. EMBO J. **2:**593–597.
20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
21. **Leary, J. J., D. J. Brigati, and D. C. Ward.** 1983. Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blots. Proc. Natl. Acad. Sci. USA **80:**4045–4049.
22. **Loveland, J., K. Gutshall, J. Kasmir, P. Prema, and J. E. Brenchley.** 1994. Characterization of psychrotrophic microorganisms with β-galactosidase activities. Appl. Environ. Microbiol. **60:**12–18.
23. **Meissner, P. S., W. P. Sisk, and M. L. Berman.** 1987. Bacteriophage λ cloning system for the construction of directional cDNA libraries. Proc. Natl. Acad. Sci. USA **84:**4171.
24. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Poch, O., H. L'Hôte, V. Dallery, F. Debeux, R. Fleer, and R. Sodoyer.** 1992. Sequence of *Kluyveromyces lactis* β-galactosidase: comparison with prokaryotic enzymes and secondary structure analysis. Gene **118:**55–63.
26. **Priest, F. J., M. Goodfellow, and C. Todd.** 1988. A numerical classification of the genus *Bacillus*. J. Gen. Microbiol. **134:**1847–1882.
27. **Schachter, H.** 1975. Enzymic microassays for D-mannose, D-glucose, D-galactose, L-fucose, and D-glucosamine. Methods Enzymol. **41:**3–10.
28. **Schmidt, B. F., R. M. Adams, C. Requadt, S. Power, and S. E. Mainzer.** 1989. Expression and nucleotide sequence of *Lactobacillus bulgaricus* β-galactosidase gene in *Escherichia coli*. J. Bacteriol. **171:**625–635.
29. **Schroeder, C. J., C. Robert, G. Lenzen, L. L. McKay, and A. Mercenier.** 1991. Analysis of the *lacZ* sequences from two *Streptococcus thermophilus* strains: comparison with the *Escherichia coli* and *Lactobacillus bulgaricus* β-galactosidase sequences. J. Gen. Microbiol. **137:**369–380.
30. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
31. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:** 4767–4771.
32. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
33. **Trimbur, D. E., K. R. Gutshall, P. Prema, and J. E. Brenchley.** 1994. Characterization of a psychrotrophic *Arthrobacter* gene and its cold-active β-galactosidase. Appl. Environ. Microbiol. **60:**4544–4552.
34. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.