Cloning and Expression of the β-d-Galactosidase Gene from *Streptococcus thermophilus* in *Escherichia coli*†

RICHARD E. HERMAN‡ AND LARRY L. MCKAY*

Department of Food Science and Nutrition, University of Minnesota, St. Paul, Minnesota 55108

Received 2 January 1986/Accepted 3 April 1986

The β-d-galactosidase (β-gal) gene from *Streptococcus thermophilus* was cloned to isolate and characterize it for potential use as a selection marker in a food-grade cloning vector. Chromosomal DNA from *S. thermophilus* 19258 was cleaved with the restriction enzyme *PstI* and ligated to pBR322 for transformation into *Escherichia coli* JM108. A β-galactosidase-positive clone was detected by its blue color on a medium supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside. This transformant possessed a single plasmid, designated pRH116, which contained, in addition to the vector DNA, a 7.0-kilobase (kb) *PstI* insertion fragment coding for β-gal activity. An extract from JM108(pRH116) contained a β-gal protein with the same electrophoretic mobility as the β-gal from *S. thermophilus* 19258. Compared with the β-gal from *E. coli* HB101, the *S. thermophilus* β-gal was of lower molecular weight. A restriction map of pRH116 was constructed from cleavage of both the plasmid and the purified insert. The construction of deletion derivatives of pRH116 with *BglII, BstEII*, and *HindIII* revealed the approximate location of the gene on the 7.0-kb fragment. The β-gal gene was further localized to a 3.85-kb region.

The recent development of gene-cloning techniques applicable to *Streptococcus lactis* (15, 17) has demonstrated the potential for the genetic engineering of dairy streptococci for the improvement of existing fermentation processes. Currently, the only cloning vectors available for these streptococci use antibiotic resistance as selective markers (7, 16, 32). Investigators have agreed that antibiotic resistance markers are unsuitable for food applications. Hence, genetic engineering of microorganisms of use in food fermentations must rely upon cloning vectors with alternative selective markers. Such vectors could conceivably incorporate carbohydrate utilization genes as selective markers.

Unlike the mesophilic dairy streptococci, *Streptococcus thermophilus* hydrolyzes lactose via a β-galactosidase (β-gal) (β-d-galactoside galactohydrolase, EC 3.2.1.23). The induction, purification, and properties of this streptococcal enzyme have been investigated (8, 25, 27, 28). Since strains of *S. thermophilus* which contained no detectable plasmids ferment lactose and exhibit β-galactosidase activity (10), the β-gal gene is assumed to be chromosomal in nature. We feel that the β-gal gene from *S. thermophilus* is a suitable candidate as a plasmid selection marker for use with food-grade organisms. The equivalent gene from *Escherichia coli* has been used extensively for cloning purposes in *E. coli* and by applying similar technology to the *S. thermophilus* β-gal gene, suitable cloning vectors can be constructed. In this study, we cloned the β-gal gene from *S. thermophilus* into *E. coli*. Restriction maps were generated, and preliminary characterization of the protein was carried out for construction of potential cloning vectors.

**MATERIALS AND METHODS**

**Bacterial strains, plasmids, and media.** *S. thermophilus* 19258 (6) was obtained from the American Type Culture Collection (Rockville, Md.). It was grown at 42°C in M17 broth (31) containing 0.5% sucrose as the fermentable carbohydrate source and was maintained by biweekly transfer.

The *E. coli* β-gal used as a reference mobility marker was obtained from cell extracts of *E. coli* HB101 (21). *E. coli* JM108 (34) was used as a recipient in all transformation experiments because of the deletion it contains Δ(lac-pro), which spans the β-gal gene (lacZ). All *E. coli* strains were grown in LB medium (21) by incubation at 37°C. When appropriate, the medium was supplemented with tetracycline (25 μg/ml), ampicillin (35 μg/ml), or chloramphenicol (34 μg/ml).

**Preparation of chromosomal DNA and plasmids.** Chromosomal DNA was isolated from *S. thermophilus* 19258 by the method of Marmur (22). Plasmid DNA was prepared by the alkaline lysis method described by Maniatis et al. (21). Purification of plasmid DNA by centrifugation to equilibrium in cesium chloride-ethidium bromide has been described before (14).

**Construction of recombinant plasmids and transformation.** The chromosomal DNA of *S. thermophilus* 19258 was cleaved with *PstI* or *EcoRI* and cloned into pBR322 (3) cleaved with *PstI* or into pSA3 (4) cleaved with *EcoRI*, respectively. The chromosomal DNA was combined with the vector DNA in a ratio of 4:1 μg at a final concentration of 50 μg/ml in ligation buffer (21) with 2 U of T4 DNA ligase (Promega BioTec, Madison, Wis.) in a final volume of 30 μl at 12°C for 16 h.

For subcloning, vector and insert DNA were combined in a picomole ends ratio of 1:2 in a 20-μl volume and a final DNA concentration of 20 μg/ml with 1 U of T4 DNA ligase and incubated at 4°C for at least 16 h. Transformation was performed as described by Davis et al. (5).

**Detection of β-gal-positive clones.** β-gal-positive clones were detected on medium supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) at 40 μg/ml (23). For *E. coli* controls, 0.2 mM isopropyl-β-d-thiogalactopyranoside was also included in the medium with X-gal. Isopropyl-β-d-
thiogalactopyranoside was not required for induction of β-gal from *S. thermophilus*.

**Agarose gel electrophoresis and restriction mapping.** Horizontal agarose gels were electrophoresed in TAE buffer (21) at 4 to 6 V/cm, stained with ethidium bromide (0.5 μg/ml), and photographed with UV transillumination. Restriction maps were constructed, and lambda DNA digested with HindIII and pBR322 digested with TaqI or HindIII were used as DNA fragment size markers (21).

Elution of DNA fragments from agarose gels. DNA fragments were separated by agarose gel electrophoresis and stained with ethidium bromide. Fragments were electroeluted onto DE-81 filter paper inserted into slits cut in front of the desired fragment. The paper was removed and placed in the barrel of a 1-ml tuberculin syringe from which the plunger and needle had been removed. The paper was washed twice by adding 200 μl of low-salt buffer (0.1 M NaCl, 0.1 mM EDTA, 10 mM Tris [pH 8.0]) in the top of the barrel and centrifuging the apparatus at 480 × g for 30 s in a centrifuge tube (15 by 100 mm). The DNA was eluted four times by the addition of 80 μl of high-salt buffer (1 M NaCl, 0.1 mM EDTA, 10 mM Tris [pH 8.0]) and precipitated with 2.5 volumes of ethanol.

**Preparation of cell extracts, polyacrylamide gel electrophoresis, and isoelectric focusing.** Cell extracts were prepared as described previously (12) from 100-ml cultures grown to late-log phase, and protein concentrations were determined by the method of Lowry et al. (19).

Polyacrylamide gel electrophoresis was performed under nondenaturing conditions with Studier-Laemmli gels (18, 30), with buffers described by Berget and Warner (1), although sodium dodecyl sulfate was not included. Samples containing 400 μg of protein were electrophoresed at a constant current of 10 mA until the tracking dye reached the bottom of the gel (about 17 h). Bands corresponding to β-gal were detected by immersing the gel in phosphate buffer containing 5 mM o-nitrophenyl-β-D-galactopyranoside at 37°C for 30 min and were photographed as described previously (9).

Nondenaturing isoelectric focusing was performed based on the procedure of O’Farrell (24) with the following modifications. To prepare the gel mixture, 1.33 ml of acrylamide/bisacrylamide (23.38/1.62%), 400 μl of amylolysates (pH 5 to 7) (40% stock solution) (Bio-Rad Laboratories, Richmond, Calif.), 100 μl of amylolysates (pH 3 to 10), and 8.17 ml of H₂O were combined, and the mixture was degassed. Polymerization was initiated by the addition of 50 μl of 10% ammonium persulfate and 5 μl of TEMED (N, N', N'-tetramethylthelylendiamine). The gel mixture was poured into tubes, overlaid with sec-butanol, and allowed to polymerize for 60 min. The sec-butanol was removed, and the gel was rinsed with water and overlaid with sample buffer (1.9 ml of H₂O, 10 μl of 0.1 M dithiothreitol, 20 μl of amylolysates [pH 3 to 10], 80 μl of amylolysates [pH 5 to 7]). The sample buffer was removed after 60 min and replaced with 1 or 2 drops of fresh sample buffer, and the tubes were filled with 0.5% ethanolamine. After the prerun, samples containing 400 μg of protein were loaded onto the gels in loading buffer (3 × loading buffer contained 0.6 ml of glycercol, 60 μl of amylolysates [pH 3 to 10], 240 μl of amylolysates [pH 5 to 7], 30 μl of 0.1 M dithiothreitol, and 1.07 ml of H₂O). Electrophoresis was performed with 0.2% H₃SO₄ in the lower chamber and 0.5% ethanolamine in the upper chamber. Detection of the β-gal in the polyacrylamide gels was as described above.

**RESULTS**

**Identification of a clone with β-gal activity.** To clone the β-gal gene, chromosomal DNA was extracted from *S. thermophilus* 19258 (a plasmid-free strain) and cleaved with *PstI* or EcoRI. Since the number of chromosomal fragments generated by cleavage with either enzyme was large, we were unable to determine the number of fragments generated. The *PstI* fragments were ligated into pBR322, and the *EcoRI* fragments were ligated into pSA3. For ligation, a standard amount of vector DNA (500 ng) was combined with an excess amount of chromosomal DNA (2 μg). These ligation mixtures were used to transform *E. coli* JM108.

None of the 43 transformants containing pSA3 was blue but 1 transformant out of 125 harboring a pBR322-derived plasmid was blue on LB medium supplemented with tetra-cycline and X-gal. This transformant, designated RH116, was Tc⁺ Ap⁺, indicating that its plasmid contained a DNA fragment in the *PstI* site. Plasmid DNA was isolated from RH116 and analyzed by comparison with pBR322 by agarose gel electrophoresis (Fig. 1). The plasmid isolated from RH116, designated pRH116, was found to be larger than pBR322, and when pRH116 was digested with *PstI*, a 7.0-kilobase (kb) insertion fragment was observed (Fig. 1). Only when this 7.0-kb *PstI* fragment was present in pBR322 could β-gal activity be detected in the isogenic host (JM108). Since JM108, with or without pBR322, forms white colonies on LB medium containing X-gal, this result indicated that the β-gal gene had been cloned into pBR322 on a 7.0-kb *PstI* fragment.

**Characterization of the pRH116-coded β-gal.** The characteristics of the cloned β-gal were compared with those in extracts of *E. coli* HB101 and the parental strain, *S. thermophilus* 19258. No β-gal activity was detected when the cell extract made from strain JM108(pBR322) was electrophoresed in a 6% polyacrylamide gel under nondenaturing conditions (Fig. 2). The extract from JM108(pRH116) demonstrated a β-gal protein having the same electrophoretic mobility as the β-gal from *S. thermophilus* 19258. The
**FIG. 2.** Locations of β-gal activity after electrophoresis of cell extracts in a 6% polyacrylamide gel. Lanes: 1, *S. thermophilus* 19258; 2, *E. coli* JM108(pBR322); 3, *E. coli* JM108(pRH116); 4, *E. coli* HB101. Enzyme activity was detected as described in the text.

Electrophoretic mobilities of these enzymes were higher than the *E. coli* β-gal from extracts of *E. coli* HB101.

Isoelectric focusing of the cell extracts indicated that both the cloned and chromosomally encoded *S. thermophilus* 19258 β-gal proteins displayed similar, if not identical, isoelectric points. The isoelectric point of the *S. thermo-

*philus* β-gal was found to be relatively neutral, whereas the *E. coli* β-gal was acidic (data not shown).

**Restriction endonuclease mapping of pRH116.** Once the clone was isolated and characterized with respect to the size of the insert and the enzyme for which it coded, a physical map of the pRH116 insert was developed by restriction enzyme analysis to allow localization of the β-gal gene on the insert and to aid in reducing the size of the fragment on which the β-gal gene is contained. The restriction map of pRH116 was constructed from cleavage of both the plasmid and the purified insert (Fig. 3). No cleavage sites were detected within the insert for the enzymes *BamHI*, *BclI*, *KpnI*, *NcoI*, *XbaI*, and *XhoI*. The enzymes *MluI* and *SalI* each had one cleavage site in the insert. The insert contained multiple cleavage sites for *BglII*, *BstEII*, *EcoRI*, *EcoRV*, *HindIII*, *PvuII*, and *SalI*. The positions of the *SalI* cleavage sites, however, are not indicated on the map.

**Localization of the β-gal gene on the 7.0-kb PstI fragment.** The restriction map of the 7.0-kb *PstI* fragment revealed that there were two cleavage sites for *BglII* and that these sites enclosed a 1.4-kb fragment. Likewise, there were two *BstEII* sites which were approximately 1.8 kb apart. Since neither *BglII* or *BstEII* has cleavage sites in pBR322 (20), these enzymes were used to construct deletion derivatives of pRH116.

To delete the 1.4-kb *BglII* fragment, pRH116 was digested with *BglII* and religated. The religation products were used to transform JM108, and transformants were selected on LB medium containing tetracycline and X-gal. All of the transformants were white except one. The plasmids were extracted from the single blue transformant and nine of the white, putative deletion derivatives. Digestion of the transformants with *BglII* revealed that the blue transformant and one of the white transformants still contained two *BglII* sites enclosing a 1.4-kb fragment; these derivatives could result

**FIG. 3.** Restriction endonuclease map of pRH116, a pBR322 derivative containing the *S. thermophilus* β-gal gene on a 7.0-kb *PstI* fragment. Numbers on the inner circle indicate the molecular size in kilobases. The zero point is the same as that of pBR322. On the outer circle, the 7.0-kb *PstI* fragment from *S. thermophilus* is indicated by the heavy line, and the vector plasmid pBR322 is indicated by the thin line. The order of the 0.98-, 0.57-, and 0.30-kb *HindIII* fragments indicated on the *PstI* insert DNA was not determined, but represents one of six possibilities.
from reinsertion or reverse reinsertion of the 1.4-kb BglII fragment, respectively. The remaining white transformants were confirmed as deletion derivatives by the presence of a single BglII site. The deletion derivative pRH120 is representative of this class (Fig. 4). The 1.8-kb BstEII fragment was deleted in a manner similar to that described for deletion of the 1.4-kb BglII fragment. The BstEII deletion derivatives, of which pRH140 is an example, were also β-gal negative (Fig. 4).

To determine if the 1.4-kb BglII fragment of pRH116 was completely responsible for β-gal activity, the purified fragment was ligated into the compatible BamHI site of pBR322 and transformed into JM108. Transformants that were Ap+ and Tc+ were white on X-gal-supplemented LB medium. These β-gal-negative transformants were shown to contain a plasmid about 1.4 kb larger than pBR322 (data not shown) and are represented by pRH130 (Fig. 4).

The 7.0-kb PsI fragment contains a 3.4-kb HindIII fragment which includes the 1.4-kb BglII fragment and the 1.8-BstEII fragment. To determine whether the β-gal gene was contained on this HindIII fragment, the purified PsI fragment was cleaved with HindIII and the products were subcloned into the HindIII site of pBR322. This resulted in the recovery of a β-gal-positive, Ap+ clone which contained only the 3.4-kb HindIII fragment in a plasmid designated pRH156 (Fig. 4). The β-gal activity of this clone was unstable, since upon subculture the clone became β-gal negative without detectable change in the 3.4-kb insert. Further attempts to obtain a β-gal-positive subclone with only the 3.4-kb HindIII fragment were unsuccessful.

Since the 3.4-kb HindIII fragment alone was insufficient for stable β-gal activity, partial HindIII digests of pRH116 were used to determine the flanking region required for enzyme activity. The minimum number of HindIII fragments required for a β-gal-positive clone was demonstrated by pRH172 (Fig. 4). This plasmid contained the 3.4-kb HindIII fragment, a 0.45-kb PsI-HindIII fragment, and a 0.3-kb HindIII-PsI fragment. The 0.45-kb fragment is the left PsI-HindIII junction fragment of the 7.0-kb PsI insert. The 0.3-kb fragment is the right HindIII-PsI junction fragment and is separated from the 3.4-kb fragment by more than 2 kb in the 7.0-kb PsI fragment. As expected, the 0.45-kb PsI-HindIII fragment alone was insufficient for β-gal activity. This fragment was recovered alone on pRH161 by the same approach used to recover pRH172 (Fig. 4).

Since it could be argued that the 0.3-kb HindIII-PsI junction fragment is involved in expression of β-gal, a further derivative, designated pRH173, was isolated (Fig. 4). This derivative was β-gal positive and did not contain the 0.3-kb HindII-PsI junction fragment. It was obtained as described for pRH172 and contained the 0.45-kb PsI-HindIII fragment and the 3.4-kb HindIII fragment. In addition, pRH173 contained two more HindIII fragments, a 0.98- and a 0.3-kb fragment. Since these fragments were absent from pRH172, they are believed to be unessential for β-gal activity. Both pRH161 and pRH173 possessed deletions into the pBR322 vector DNA (see Fig. 3 for HindIII site on pBR322).

**DISCUSSION**

The hydrolysis of lactose to its component monosaccharides (glucose and galactose) by β-gal is of interest for several reasons. First, lactose is digested with difficulty by a...
large proportion of the world population (11). Also, lactose has low solubility in water, which leads to problems in the concentration of whey and in the preparation of certain food items (27). Additionally, lactose has a relatively low level of sweetness. These problems can be overcome, to a large extent, by hydrolysis of lactose to its monosaccharides which are sweeter, more soluble, and more digestible than lactose (27). Investigators have screened for microorganisms capable of producing high levels of β-gal (2, 26, 29, 33). Although a number of sources for β-gal now exist, Greenberg and Mahoney have suggested that S. thermophilus is a promising source for the production of this enzyme because it is a food-approved organism and also because S. thermophilus β-gal is more heat stable than β-gal from other sources (8). Therefore, the cloning of this enzyme may have commercial value, in addition to its potential use as a selection marker in a food-grade cloning vector. The cloning of the S. thermophilus β-gal gene would also facilitate the study of its regulation and thereby improve our understanding of lactose metabolism in this industrially important bacterium.

The gene for β-gal in S. thermophilus was assumed to be chromosomally mediated, since plasmid-free strains ferment lactose and form blue colonies on medium containing X-gal. The size of the S. thermophilus chromosome was assumed to be similar to that reported for group N streptococci, which is approximately 2 x 10⁶ daltons or 3 x 10⁵ kb pairs (13). The assumption that most plasmid-cloning vectors can accommodate only about 10 kb of inserted DNA indicated that at least 300 clones were needed to establish a chromosomal gene bank. This dictated that the cloning be conducted in E. coli, since the techniques for transformation are very efficient as compared with S. lactis. It was also assumed that the β-gal gene from S. thermophilus would be expressed in E. coli.

By using pBR322, DNA extracted from S. thermophilus 19258 was shotgun cloned into the E. coli host JM108, resulting in a β-gal-positive clone designated JM108 (pRH116). Psrl digestion showed that, in addition to the vector DNA, pRH116 contained a 7.0-kb insertion. The same electrophoretic mobility and isoelectric point as the enzyme in extracts from S. thermophilus was also present in JM108(pRH116) extracts was shown to have the same electrophoretic mobility and isoelectric point as the enzyme in extracts from S. thermophilus 19258 and was clearly distinguishable from the E. coli β-gal in extracts from HB101. When JM108 contained only pBR322, no β-gal activity was detected in cell extracts or by plating on LB medium containing X-gal. These data support the conclusion that the S. thermophilus β-gal gene had been cloned on a 7.0-kb Psrl fragment and expressed in E. coli.

The β-gal of S. thermophilus 19258 has been shown by sodium dodeyl sulfate-polyacrylamide gel electrophoresis to have a molecular weight of 105,000 (28). The size of the protein indicates that approximately 2.3 kb of DNA would be required to code for this β-galactosidase. Since the cloned β-gal was contained on a 7.0-kb fragment, we sought to obtain a subclone with the minimal amount of DNA required to code for the S. thermophilus β-gal gene. The construction of β-gal-negative deletion derivatives of pRH116 with BglII and BsrEII revealed the approximate location of the gene on the 7.0-kb fragment. Attempts to obtain an EcoRI subclone with enzyme activity were unsuccessful, probably because of the high frequency of EcoRI cleavage sites in the approximate location of the β-gal gene. By partial digestion of pRH116 with HindIII, we were able to obtain a β-gal-positive clone contained on a 4.15-kb fragment, which consisted of a 0.45-kb Psrl-HindIII, a 3.4-kb HindIII, and a 0.3-kb HindIII-Psrl fragment. A subclone containing only the 0.45-kb fragment was insufficient for β-gal activity. A subclone containing only the 3.4-kb HindIII fragment was initially isolated as a β-gal-positive clone based on the formation of a blue colony on X-gal plating medium, but it became negative upon subculturing. We cannot explain this observation, but the possibility that both pRH156 and pRH157 became β-gal negative as the result of an undetectable deletion cannot be discounted. This result suggests that the β-gal gene is located on a portion of the 3.4-kb HindIII fragment and that the 0.45-kb Psrl-HindIII fragment which adjoins it is required for stable expression of the β-gal gene. Since the 0.45-kb fragment was contiguous with the 3.4-kb fragment in the original insert, and the 0.3-kb fragment was not, we assumed that the 0.45- and 3.4-kb fragments together were necessary for β-gal activity. This was substantiated by the isolation of a β-gal-positive subclone harboring pRH173 which contained the 0.45- and the 3.4-kb fragments, but not the 0.3-kb fragment. Therefore, the β-gal gene was localized to a 3.85-kb region. Further studies will be required to determine the exact location of the β-gal gene and whether the 7.0-kb Psrl fragment codes for other proteins involved in lactose utilization by S. thermophilus.

ACKNOWLEDGMENTS

We appreciate the use of D. Peter Snustad's laboratory and the technical assistance of Nancy Haas for the isoelectric focusing experiments. We also acknowledge the technical assistance of Nancy Maxfield.

This research was supported in part by General Mills, Inc., Minneapolis, Minn., and was conducted in part under Minnesota Agricultural Experiment Station Project 18-62, supported by Hatch and General Agricultural Research funds.

LITERATURE CITED