Proline biosynthesis in *Streptococcus thermophilus*: characterization of the proBA operon and its products

Danila Limauro, Angela Falciatore, Anna L. Basso, Giuseppe Forlani and Maurilio De Felice

The presence of proline in the medium was not essential for growth of *Streptococcus thermophilus*, indicating that there is a proline biosynthetic pathway in this organism. Genetic and biochemical analysis identified and characterized this pathway. Two genes, designated proB and proA, were cloned, sequenced and characterized. Biochemical analysis of the proB- and proA-encoded enzymes showed that the proline biosynthetic pathway of *S. thermophilus* is similar to the one previously described in *Escherichia coli*. The deduced amino acid sequence of a 2408 kb DNA region containing the genes revealed the similarity of the *S. thermophilus* gene products to ProB and ProA of *E. coli* and *Serratia marcescens*, and to the corresponding N- and C-terminal domains of the bifunctional plant enzyme Δ¹-pyrroline-5-carboxylate synthetase of *Vigna aconitifolia*. Northern blot analysis showed that the two genes in *S. thermophilus* are organized in a single operon with proB proximal and proA distal to the promoter; primer extension analysis indicated that proBA transcription is not under repressive control by exogenously supplied proline.

**Keywords**: proline biosynthesis, *Streptococcus thermophilus*, gene structure and expression, γ-glutamyl kinase, γ-glutamyl phosphate reductase

**INTRODUCTION**

Knowledge of the molecular genetics of lactic acid bacteria has progressed impressively during the last few years, mostly because of their relevance in biotechnology, and a number of genes important for industrial applications have been characterized, especially in *Lactococcus* spp. More limited attention has been paid to basic physiological features of these organisms, knowledge of which is essential to improve and expand industrial applications.

Because of their nutritional fastidiousness, lactic acid bacteria are unable to grow unless certain amino acids are supplied in the medium. This differentiates them from bacteria such as *Escherichia coli* and *Bacillus subtilis*, which can grow in media containing only mineral salts and glucose. The recent characterization of tryptophan, histidine and branched-chain amino acid biosynthesis operons in lactococcal strains (Bardowski et al., 1992; Delorme et al., 1992; Godon et al., 1992) has provided some information on gene organization and regulation. A comparison of equivalent genes derived from different strains reveals that inactive amino acid biosynthetic genes are present in lactococcal dairy starters (Delorme et al., 1993).

Despite its importance in the dairy industry (especially for the production of yoghurt and many Italian, French and Swiss cheeses), *Streptococcus thermophilus* has been the subject of very limited physiological and genetic studies. Based on procedures with better-known lactic acid bacteria, growth media are normally supplemented with various amino acids, for which a growth requirement is in some cases only putative and the biosynthesis of which has never been investigated. Among these is proline, an important amino acid which is the most abundant in casein. It stimulates cell growth in some lactic acid bacteria (Smid & Konings, 1990) and may be involved in osmoregulation, as has been proven for many microorganisms (Bae & Miller, 1992; Beumer et al., 1994;
Fig. 1. Metabolic pathways for proline biosynthesis. ProB, ProA and ProC are the biosynthetic enzymes defined by E. coli structural genes proB, proA and proC. Broken and thick arrows indicate reactions occurring specifically in B. subtilis and C. sporogenes, respectively.


Proline biosynthesis has been studied in E. coli and other Enterobacteriaceae as well as in Bacillus and Streptomyces spp. (Baumberg & Klingel, 1993; Campanile et al., 1993; Fons, 1987), and has been reported to arise through the pathway shown in Fig. 1. Three enzymes, γ-glutamyl kinase (EC 2.7.2.11), γ-glutamylphosphate reductase (EC 1.2.1.41) and Δ1-pyrroline-5-carboxylate (PSC) reductase (EC 1.5.1.2), encoded in E. coli by the proB, proA and proC genes, respectively, are responsible for the conversion of glutamic acid to proline (Leisinger, 1987). In Clostridium sporogenes, ornithine is directly converted to proline by the action of ornithine cyclase (Fons et al., 1987), while in B. subtilis proline may also be synthesized through the formation of PSC from ornithine, catalysed by ornithine transaminase (Baumberg & Klingel, 1993; Fons et al., 1987).

In this article, we show that the wild-type S. thermophilus strain ATCC 19258 is able to grow in proline-free medium. We describe the isolation and characterization of the proBA operon and its gene products γ-glutamyl kinase and γ-glutamylphosphate reductase, which allowed us to identify the proline biosynthetic pathway of this organism.

METHODS

Bacterial strains and plasmids. The following strains were used in this study: E. coli DH5α (Sambrook et al., 1989); ID307 [proΔ(BA)] (laboratory collection); χ342 (proC29) (B. Bachmann, E. coli Genetic Stock Center, Yale University, CT, USA); S. thermophilus ATCC 19258.

All plasmids constructed were derived from pUC18 and pUC19 vectors of E. coli (Yanisch-Perron et al., 1985) in which S. thermophilus DNA was recombined. Plasmid pST74a contains a partial genomic 5.7 kb Sau3AI fragment in pUC19; pST74b contains the same insert digested with restriction endonucleases SmaI and SphI, recombined in pUC18. The insert was reduced in size by digestion with restriction endonucleases in pST75, pST84 and pST79 and by unidirectional deletion with the Erase-a-base system (Promega) in pST80. Plasmid pBS2 is a pACYC177 derivative with a proBA insert of Streptococcus bovis (Campanile et al., 1993).

Growth media. Luria-Bertani (LB) broth (Miller, 1972) and SOB broth (Hanahan, 1983) were used for cultivating E. coli in rich media; M9 minimal medium (Miller, 1972) containing 0.5% (w/v) glucose and, where required, 50 μg ampicillin ml−1, 0.4 mM methionine and 0.4 mM proline, was used for complementation tests. S. thermophilus strain ATCC 19258 was grown at 42°C in M17 medium (Terzaghi & Sandline, 1975) with 1% (w/v) glucose. The chemically defined medium (JDR) used (Reiter & Oram, 1962) was modified as described previously by Molenaar et al. (1993).

Chemicals. DL-PSC was synthesized and purified as previously described by Williams & Frank (1975). Sephacryl S200, Sephacryl S300HR and M₄ markers for column calibration were purchased from Pharmacia. All other biochemicals were from Sigma.

Preparation of cell-free extracts and gel-filtration chromatography. Crude extracts were prepared as previously described by Campanile et al. (1993) and treated with solid ammonium sulfate (70%, w/v, saturation); the precipitated proteins were collected by centrifugation at 12000 g for 20 min, resuspended in extraction buffer and further centrifuged as above. Extracts were then diluted to a concentration of 5 mg protein ml⁻¹ and subjected to gel-filtration chromatography. Aliquots (2 ml) were layered onto either a Sephacryl S200 (87 x 16 cm) or a Sephacryl S300 HR (82 x 16 cm) column equilibrated in extraction buffer. Elution proceeded at a constant flow of 10 ml h⁻¹, collecting one 20 ml fraction and 90 1 ml fractions. M₄ markers for column calibration were: bovine thyroid thyroglobulin (M₄ 669000), horse spleen ferritin (M₄ 440000), bovine liver catalase (M₄ 232000), rabbit muscle aldolase (M₄ 158000) bovine serum albumin (M₄ 67000), hen egg ovalbumin (M₄ 43000), bovine pancreas chymotrypsinogen A (M₄ 25000) and bovine pancreas ribonuclease A (M₄ 13700).

Enzyme assays. γ-Glutamyl kinase and γ-glutamylphosphate reductase activities were measured as described previously by Campanile et al. (1993). The activity of the kinase was assayed by measuring the production of γ-glutamyl hydroxamate from glutamate and ATP; since the reaction is also catalysed by glutamine synthetase (EC 6.3.1.2), parallel assays including 100 mM l-proline as a specific feedback inhibitor of γ-glutamyl kinase were also performed. One unit of enzyme activity was defined as the proline-inhibited production of 1 nmol hydroxamate s⁻¹. The activity of the reductase was assayed in the reverse direction by measuring the formation of NADPH from NADP⁺, P₄ and PSC; as the same reaction could be catalysed also by PSC dehydrogenase (EC 1.5.1.12), which does not require phosphate, parallel assays in the absence of P₄ were also performed. One unit of enzyme activity was defined as the phosphate-dependent reduction of 1 nmol NADP⁺ s⁻¹. Protein concentration was measured by the method of Bradford (1976) using bovine serum albumin as a standard.
DNA manipulations and molecular cloning. *S. thermophilus* chromosomal DNA was prepared as previously described by Slos et al. (1991). Southern blotting and DNA hybridization were performed according to standard procedures (Sambrook et al., 1989); [α-32P]dCTP-labelled DNA probes were prepared by random primer labelling (Amersham). The filters (Amersham) were washed under stringent conditions [1 x SSC (0.15 M sodium chloride plus 0.015 M sodium citrate)/0.1 %, w/v, SDS at 55 °C]. A genomic bank of 5-8 kb fragments partially digested with Sac3AI was purified on a sucrose gradient (Hopwood et al., 1985) and cloned into pUC19 (BamHI site). The ligation mixture was used to transform *E. coli* DH5α (Hanahan, 1983) and the library was screened by colony hybridization (Sambrook et al., 1989); clones that gave positive signals were purified and their plasmid DNA was extracted and used to transform *E. coli* proline auxotrophic strains as previously described by Sambrook et al. (1989).

DNA sequence analysis. DNA was sequenced by the dideoxy chain-termination method (Sanger et al., 1977) using the Sequenase system (United States Biochemical).

Assembling and editing of the sequencing data, analysis of the putative ORFs and detection of restriction sites were performed with DNA strider software by the Institut de Recherche Fondamentale, Commissariat à l'Energie Atomique, Paris, France. The Gene Works (Intelligenetics) program was used to screen EMBL/GenBank databases for sequence identity and alignment with homologous sequences.

RNA isolation and analysis. Total RNA was extracted from 50 ml exponentially growing cultures of *S. thermophilus* as previously described by Cutting et al. (1989). The primer extension analysis was performed as reported by Igo & Losick (1986) using two synthetic 20-base oligonucleotides, 5'-CCCTTACGCCAGCGTGGGAA-3' (MX19) and 5'-CCCTTACGCCAGCGTGGGAA-3' (NO12), complementary to nucleotides 233-252 and 259-278, respectively (Fig. 4). The products of primer extension were analysed on 6% (w/v) polyacrylamide gels.

Northern blot analysis. RNA extracted from exponentially growing cells was analysed on a 1% (w/v) agarose gel containing formaldehyde according to standard protocols (Sambrook et al., 1989). RNA was transferred to Hybond (Amersham) membranes and hybridized with a [α-32P]dCTP-labelled BglII-BglII DNA fragment (877 bp) as described by Thomas (1980).

RESULTS

Growth of *S. thermophilus* in proline-free medium

Cultures of *S. thermophilus* strain ATCC 19258 grew exponentially in JDR medium, regardless of the presence of proline, but the generation time was somewhat higher in the absence than in presence of proline (162 and 130 min, respectively). This suggested that proline biosynthesis does occur in *S. thermophilus*, although the amount of product might not be optimal for protein synthesis during exponential growth.

Cloning of the proBA region

To identify the genetic information for proline biosynthesis and to characterize the corresponding pathway, total DNA extracted from *S. thermophilus* ATCC 19258 was partially digested with Sac3AI and purified on a sucrose gradient. A genomic library was constructed in pUC19 linearized with BamHI. The ligation mixture was used to transform *E. coli* DH5α; recombinant clones showing resistance to ampicillin (AmpR) were selected on LB agar containing X-Gal/IPTG. White AmpR clones were screened in situ hybridization using the insert of plasmid pSB2 containing the proBA region of *S. bovis* (Campanile et al., 1993) as a probe, which was found in a preliminary Southern blot experiment to hybridize with discrete bands of *S. thermophilus* DNA. Five positive clones were obtained.

Recombinant plasmids from these clones were extracted and used to transform the *E. coli* auxotrophic strain ID307 (ΔproBA). One of the plasmids, capable of complementing the ΔproBA mutation of strain ID307, pST74a (Fig. 2), was isolated and chosen for further analysis. The 5.7 kb insert was characterized by restriction analysis, and cloned in the opposite orientation in pUC18 (yielding plasmid pST74b). Several individual AmpR clones were selected on LB medium and tested for growth on minimal medium. All of those tested were prototrophic, indicating that putative pro genes were expressed from an internal promoter.

Fig. 2 shows a relevant physical map of pST74a and some plasmids derived from it, as well as the Pro+ phenotype of *E. coli* transformants. The smallest fragment able to complement the ΔproBA mutation of strain ID307 was the 3.25 kb insert of pST80. None of the plasmids complemented the proC mutation in strain χ342.

Biochemical evidence of γ-glutamyl kinase and γ-glutamylphosphate reductase activities in *S. thermophilus*

To check whether pST74a encoded γ-glutamyl kinase and γ-glutamylphosphate reductase, these activities were assayed in crude extracts of *E. coli* strain ID307(pST74a) and, as a control, ID307(pUC19). The eluate from a Sephacryl S300HR column for γ-glutamyl kinase activity showed two peaks in ID307(pST74a) extracts, while only one was detectable in extracts from cells transformed with the vector alone (Fig. 3a); the presence of a γ-glutamyl kinase inhibited by proline in an *E. coli* (ΔproBA) strain has been reported previously (Campanile et al., 1993; Hayzer & Moses, 1978). Following fractionation of extracts from ID307(pST74a) cells on a Sephacryl S200 column, a single peak of γ-glutamylphosphate reductase was found (Fig. 3b), while no activity was detected in the extracts from cells transformed with the vector alone. Native Mr values of approximately 240000 and 180000 for γ-glutamyl kinase and γ-glutamylphosphate reductase, respectively, were measured by gel filtration (data not shown).

Nucleotide sequence of the proBA region

The nucleotide sequence of the 2-408 kb region from the left-hand *ArrII* site to the *XbaI* site of plasmid pST80 was determined (Fig. 4). Analysis of this sequence for potential protein coding regions revealed two tandem ORFs. The
first extended from an ATG at position 141 to a TAA termination codon at position 944 and consisted of 801 nucleotides; the predicted start site was not preceded by a good consensus sequence for a ribosome-binding site (RBS). The second ORF started one nucleotide downstream of the termination codon of the first ORF, was preceded by a typical RBS complementary to the 3'-end of the lactococcal 16S RNA (UCCUCC) (de Vos, 1987) and terminated with a TAA codon at position 2196. A rho-independent transcription terminator-like structure was present downstream of the putative proB gene.

The two ORFs were compared with sequences in the EMBL, GenBank and SWISS-PROT databases using the Gene Works program. The ORF of 267 amino acids showed an overall identity of 26.9% and 27.7% with ProB of *E. coli* (Deutch et al., 1984) and *Serratia marcescens* (Omori et al., 1991), respectively, and an overall identity of 31.6% with the N-terminal region of P5C synthetase of *Vigna aconitifolia*, a bifunctional enzyme that catalyses the first two steps of proline biosynthesis in plants (Hu et al., 1992). The ORF of 416 amino acids showed an overall identity of 45.3%, 47.9% and 31.8% with ProA of *E. coli* and *Ser. marcescens*, and with the C-terminal region of P5C synthetase of *V. aconitifolia*, respectively. Homology was lower when the two *S. thermophilus* ORFs were compared with all other known proline biosynthetic genes.

**RNA analysis**

To verify that the proB and proA genes form an operon, we prepared RNA from cells grown in JDR medium. RNA was analysed by Northern blotting, using an 877 bp BgIII-BgIII fragment that contains the 3'-end of proB and the 5'-portion of proA (Fig. 2) as a probe. Only one transcript of about 2.1 kb was detected (not shown).

A primer extension analysis was used to identify the transcription start point of proB mRNA. Oligonucleotide MX19 (see Methods), labelled at its 5'-end, was designed to anneal to the proximal part of proB mRNA (Fig. 4). The 5'-end, as defined by the band corresponding to the primer-extended product (Fig. 5), was identical with the first nucleotide (nucleotide 141) of the proB translation initiation codon AUG. The intensity of the band was the same for RNA extracted from cells grown in proline-containing (Fig. 5, lane 1) and proline-free (Fig. 5, lane 2) JDR medium, indicating that the presence of the amino acid does not repress proB transcription.

To confirm these results, we repeated the primer extension
analysis with a different oligonucleotide, complementary to nucleotides 259–278 (N012; see Methods); the results agreed with those found previously (data not shown).

A possible promoter (ATGT-TATAAT) (Fig. 4) was present upstream of the transcription start site, and showed a good consensus to the canonical sequence of

Fig. 4. Nucleotide and deduced amino acid sequences of the *S. thermophilus* proBA region. Gene names are indicated at the beginning of each amino acid sequence. Putative -35 and -10 boxes of the promoter, RBS of proA and a possible transcription terminator are underlined. Asterisks indicate proB and proA stop codons.

...
auxotroph, although the growth rate was enhanced by addition of proline to the medium. We report the first transcribed nucleotide. Vos, 1987), principally in the various other amino acids, proved not to be a proline alongside the dideoxy sequencing products primed by the same primer on the coding strand for size comparison. Lanes: 1 and 2, primer-extension product with RNA extracted from primer extension analysis. The extended products are shown extended products of 110 bp obtained using oligonucleotide MX19 (see Methods); the arrowhead on the right indicates the standard wild-type strain of S. thermophilus E. coli, ATCC 19258), normally grown in media containing proline and various other amino acids, calculated on the basis of the predicted amino acid sequences, were 28947 and 45361, respectively; these values, compared with the \( M_r \) detected by gel filtration (243000 and 180000 for \( \gamma \)-glutamyl kinase and \( \gamma \)-glutamylphosphate reductase, respectively) suggest an octameric composition for \( \gamma \)-glutamyl kinase and a tetrameric composition for \( \gamma \)-glutamylphosphate reductase. The enzymes of E. coli are a hexamer and an octamer, respectively, of identical subunits (Leisinger, 1987).

**DISCUSSION**

A standard wild-type strain of S. thermophilus (ATCC 19258), normally grown in media containing proline and various other amino acids, proved not to be a proline auxotroph, although the growth rate was enhanced by addition of proline to the medium. We report the molecular cloning and physical characterization of a DNA fragment, biochemical analysis of which identified the proline biosynthetic pathway. The fragment encodes \( \gamma \)-glutamyl kinase and \( \gamma \)-glutamylphosphate reductase, which are the products of \( proB \) and \( proA \), respectively, in E. coli (Deutch et al., 1984) and Ser. marcescens (Omori et al., 1991). Physical and biochemical analysis of the genes and their products allowed us to establish that the proline biosynthetic pathway of S. thermophilus resembles that described previously in some other bacteria (Baumberg & Klingel, 1993).

As previously observed in E. coli (Deutch et al., 1984; Leisinger, 1987), the two genes constitute an operon with a single promoter proximal to \( proB \). The transcriptional start, as detected by primer extension analysis, was coincident with the A of the putative first AUG codon of \( proB \) and was not preceded by a canonical RBS (Fig. 4). Absence of untranslated mRNA at the 5'-end, although not usual, has been reported in several prokaryotic transcription units in both eubacteria and archaea (Stormo et al., 1982). These include transcripts derived from the pRM promoter of \( \lambda \) phage (Pottage et al., 1976), the tetE gene of transposon Tn1721 (Klock & Hillen, 1986), the \( ermA \) gene of Streptomyces erythraeus (Bibb et al., 1986), the sta gene of Streptomyces lavendulae (Horinouchi et al., 1987), the apb gene of Streptomyces fradiae (Janssen et al., 1989) and the hisD gene of Streptomyces coelicolor (Limauro et al., 1992). Since a canonical RBS is missing, it is possible that interaction of mRNA with 16S rRNA for \( proB \) translation involves nucleotides within the 5'-end of the coding region. This is the case for E. coli, where the so-called downstream box, a sequence of approximately 15 nucleotides located just downstream of the first AUG codon, promotes efficient transcription in the absence of a canonical RBS, because it interacts with a 3'-portion of 16S rRNA not far from the RBS, the so-called anti-downstream box (Sprengart et al., 1996). In this connection, it is interesting to note that there is a high degree of homology (only two mismatches within an 11-nucleotide stretch) between nucleotides 153-164 in Fig. 4 and nucleotides 1467-1478 of S. thermophilus 16S rRNA (Ludwig et al., 1992).

The homology of the putative sequences of S. thermophilus gene products with ProB and ProA of E. coli and Ser. marcescens was higher for ProA than for ProB; this is not surprising in view of the fact that the \( proB \) gene in S. thermophilus is shorter than in the other two bacteria (267 vs 367 amino acids), which suggests a distinctive evolutionary route.

**Fig. 5.** Mapping of the 5'-end of mRNA from the \( proBA \) operon by primer extension analysis. The extended products are shown alongside the dideoxy sequencing products primed by the same primer on the coding strand for size comparison. Lanes: 1 and 2, primer-extension product with RNA extracted from S. thermophilus grown in JDR medium with and without 50 \( \mu \)g proline ml\(^{-1}\), respectively; T, C, G, A, DNA sequencing tracks. The arrowhead on the left of the figure indicates primer-extended products of 110 bp obtained using oligonucleotide MX19 (see Methods); the arrowhead on the right indicates the first transcribed nucleotide.
Upstream of the proBA transcriptional start point detected by primer extension analysis is a sequence resembling the canonical sequence of lactococcal promoters (Fig. 4). We could not detect a repressive effect of exogenous proline on proBA transcription (Fig. 5), possibly indicating that, in S. thermophilus as in E. coli and Salmonella typhimurium (Brady & Csonka, 1988), there is no negative end-product control of proBA transcription. Thus proline biosynthesis in S. thermophilus may occur through low-level constitutive expression of proline biosynthetic enzymes. This may not supply sufficient proline for growth in chemically defined media at 42°C; the resulting de facto starvation for proline would explain why addition of the amino acid to the growth medium enhances the growth rate. This feature would not limit S. thermophilus metabolism in natural habitats because (i) growth rate is influenced by many metabolic factors, and (ii) some supply of proline derives from proteolysis of casein.

ACKNOWLEDGEMENTS

We thank E. Ricca for many helpful suggestions and critical reading of the manuscript, R. Frunzio and M. Varcamonti for reading of the manuscript, R. Frunzio and M. Varcamonti for many helpful suggestions and critical reading of the manuscript, R. Frunzio and M. Varcamonti for partial support was also obtained from the Piano Nazionale Biotecnologie Vegetali of the MIRAAF.

REFERENCES


*p*3282

Received 4 April 1996; revised 9 July 1996; accepted 17 July 1996.