The gene encoding IIAB\textsuperscript{Man} in \textit{Streptococcus salivarius} is part of a tetracistronic operon encoding a phosphoenolpyruvate:mannose/glucose phosphotransferase system

Louis-André Lortie, Michel Pelletier, Christian Vadeboncoeur and Michel Frenette

Author for correspondence: Michel Frenette. Tel: +1 418 656 2131 ext. 5502. Fax: +1 418 656 2861. e-mail: Michel.Frenette@rebb.ulaval.ca

INTRODUCTION

Oral streptococci are among the most abundant micro-
organisms in the oral environment (Marsh & Martin, 1992). These Gram-positive eubacteria use sugars as their main source of energy, and transport several of them by the phosphoenolpyruvate:sugar phosphotransferase system (PTS) (Vadeboncoeur & Pelletier, 1997). The PTS concomitantly transports and phosphorylates sugars at the expense of phosphoenolpyruvate.

Glucose and mannose are transported in streptococci by the mannose-PTS (phosphoenolpyruvate:mannose phosphotransferase system), which consists of a cytoplasmic IIA protein, called IIAB\textsuperscript{Man}, and an uncharacterized membrane permease. This paper reports the characterization of the \textit{man} operon encoding the specific components of the mannose-PTS of \textit{Streptococcus salivarius}. The \textit{man} operon was composed of four genes, \textit{manL}, \textit{manM}, \textit{manN} and \textit{manO}. These genes were transcribed from a canonical promoter (P\textit{man}) into a 3.6 kb polycistronic mRNA that contained a 5'-UTR (untranslated region). The predicted \textit{manL} gene product encoded a 35.5 kDa protein and contained the amino acid sequences of the IIA and IIB phosphorylation sites already determined from purified \textit{S. salivarius} IIAB\textsuperscript{Man}. Expression of \textit{manL} in \textit{Escherichia coli} generated a 35 kDa protein that reacted with anti-IIAB\textsuperscript{Man} antibodies. The predicted ManM protein had an estimated size of 27.2 kDa. ManM had similarity with IIC domains of the mannose-EII family, but did not possess the signature proposed for mannose-IIC proteins from Gram-negative bacteria. From multiple alignment analyses of sequences available in current databases, the following modified IIC\textsuperscript{Man} signature is proposed: GX,G(DNH)X,G(LIVM)X,G(SSL)[LT][EQ]. The deduced product of \textit{manN} was a hydrophobic protein with a predicted molecular mass of 33.4 kDa. The ManN protein contained an amino acid sequence similar to the signature sequence of the IID domains of the mannose-EII family. \textit{manO} encoded a 13.7 kDa protein. This gene was also transcribed as a monocistronic mRNA from a promoter located in the \textit{manN-manO} intergenic region. A search of current databases revealed the presence of IIAB\textsuperscript{Man}, ManM, ManN and ManO orthologues in \textit{Streptococcus mutans}, \textit{Streptococcus pyogenes}, \textit{Streptococcus pneumoniae} and \textit{Enterococcus faecalis}. This work has elucidated the molecular structure of the mannose PTS in streptococci and enterococci, and demonstrated the presence of a putative regulatory protein (ManO) within the \textit{man} operon.

Keywords: enzyme II, streptococci, \textit{man} operon, phosphotransferase system

INTRODUCTION

Oral streptococci are among the most abundant micro-
organisms in the oral environment (Marsh & Martin, 1992). These Gram-positive eubacteria use sugars as their main source of energy, and transport several of them by the phosphoenolpyruvate:sugar phosphotransferase system (PTS) (Vadeboncoeur & Pelletier, 1997). The PTS concomitantly transports and phosphorylates sugars at the expense of phosphoenolpyruvate.

Glucose and mannose are transported in streptococci by the mannose-PTS (phosphoenolpyruvate:mannose phosphotransferase system), which consists of a cytoplasmic IIA protein, called IIAB\textsuperscript{Man}, and an uncharacterized membrane permease. This paper reports the characterization of the \textit{man} operon encoding the specific components of the mannose-PTS of \textit{Streptococcus salivarius}. The \textit{man} operon was composed of four genes, \textit{manL}, \textit{manM}, \textit{manN} and \textit{manO}. These genes were transcribed from a canonical promoter (P\textit{man}) into a 3.6 kb polycistronic mRNA that contained a 5'-UTR (untranslated region). The predicted \textit{manL} gene product encoded a 35.5 kDa protein and contained the amino acid sequences of the IIA and IIB phosphorylation sites already determined from purified \textit{S. salivarius} IIAB\textsuperscript{Man}. Expression of \textit{manL} in \textit{Escherichia coli} generated a 35 kDa protein that reacted with anti-IIAB\textsuperscript{Man} antibodies. The predicted ManM protein had an estimated size of 27.2 kDa. ManM had similarity with IIC domains of the mannose-EII family, but did not possess the signature proposed for mannose-IIC proteins from Gram-negative bacteria. From multiple alignment analyses of sequences available in current databases, the following modified IIC\textsuperscript{Man} signature is proposed: GX,G(DNH)X,G(LIVM)X,G(SSL)[LT][EQ]. The deduced product of \textit{manN} was a hydrophobic protein with a predicted molecular mass of 33.4 kDa. The ManN protein contained an amino acid sequence similar to the signature sequence of the IID domains of the mannose-EII family. \textit{manO} encoded a 13.7 kDa protein. This gene was also transcribed as a monocistronic mRNA from a promoter located in the \textit{manN-manO} intergenic region. A search of current databases revealed the presence of IIAB\textsuperscript{Man}, ManM, ManN and ManO orthologues in \textit{Streptococcus mutans}, \textit{Streptococcus pyogenes}, \textit{Streptococcus pneumoniae} and \textit{Enterococcus faecalis}. This work has elucidated the molecular structure of the mannose PTS in streptococci and enterococci, and demonstrated the presence of a putative regulatory protein (ManO) within the \textit{man} operon.

Keywords: enzyme II, streptococci, \textit{man} operon, phosphotransferase system

INTRODUCTION

Oral streptococci are among the most abundant micro-
organisms in the oral environment (Marsh & Martin, 1992). These Gram-positive eubacteria use sugars as their main source of energy, and transport several of them by the phosphoenolpyruvate:sugar phosphotransferase system (PTS) (Vadeboncoeur & Pelletier, 1997). The PTS concomitantly transports and phosphorylates sugars at the expense of phosphoenolpyruvate.

Glucose and mannose are transported in streptococci by the mannose-PTS (phosphoenolpyruvate:mannose phosphotransferase system), which consists of a cytoplasmic IIA protein, called IIAB\textsuperscript{Man}, and an uncharacterized membrane permease. This paper reports the characterization of the \textit{man} operon encoding the specific components of the mannose-PTS of \textit{Streptococcus salivarius}. The \textit{man} operon was composed of four genes, \textit{manL}, \textit{manM}, \textit{manN} and \textit{manO}. These genes were transcribed from a canonical promoter (P\textit{man}) into a 3.6 kb polycistronic mRNA that contained a 5'-UTR (untranslated region). The predicted \textit{manL} gene product encoded a 35.5 kDa protein and contained the amino acid sequences of the IIA and IIB phosphorylation sites already determined from purified \textit{S. salivarius} IIAB\textsuperscript{Man}. Expression of \textit{manL} in \textit{Escherichia coli} generated a 35 kDa protein that reacted with anti-IIAB\textsuperscript{Man} antibodies. The predicted ManM protein had an estimated size of 27.2 kDa. ManM had similarity with IIC domains of the mannose-EII family, but did not possess the signature proposed for mannose-IIC proteins from Gram-negative bacteria. From multiple alignment analyses of sequences available in current databases, the following modified IIC\textsuperscript{Man} signature is proposed: GX,G(DNH)X,G(LIVM)X,G(SSL)[LT][EQ]. The deduced product of \textit{manN} was a hydrophobic protein with a predicted molecular mass of 33.4 kDa. The ManN protein contained an amino acid sequence similar to the signature sequence of the IID domains of the mannose-EII family. \textit{manO} encoded a 13.7 kDa protein. This gene was also transcribed as a monocistronic mRNA from a promoter located in the \textit{manN-manO} intergenic region. A search of current databases revealed the presence of IIAB\textsuperscript{Man}, ManM, ManN and ManO orthologues in \textit{Streptococcus mutans}, \textit{Streptococcus pyogenes}, \textit{Streptococcus pneumoniae} and \textit{Enterococcus faecalis}. This work has elucidated the molecular structure of the mannose PTS in streptococci and enterococci, and demonstrated the presence of a putative regulatory protein (ManO) within the \textit{man} operon.

Keywords: enzyme II, streptococci, \textit{man} operon, phosphotransferase system

INTRODUCTION

Oral streptococci are among the most abundant micro-
organisms in the oral environment (Marsh & Martin, 1992). These Gram-positive eubacteria use sugars as their main source of energy, and transport several of them by the phosphoenolpyruvate:sugar phosphotransferase system (PTS) (Vadeboncoeur & Pelletier, 1997). The PTS concomitantly transports and phosphorylates sugars at the expense of phosphoenolpyruvate.
pyruvate (PEP). It is composed of the general proteins HPPr and Enzyme I (EI), which are found in the cytoplasm, as well as the sugar-specific proteins called Enzymes II (EII). EI catalyses the transfer of a phosphate group from PEP to HPPr. HPPr is a phosphate carrier that shuttles between EI and the EII complexes. These complexes are composed of three or occasionally four domains (IIA, IIB, IIC and IID) that can be located on one or multiple polypeptide chains. The IIA domain contains a histidine residue that is phosphorylated by HPPr → P. IIA → P transfers its phosphate group to the IIB domain, usually on a cysteine, except for the IIB proteins of the mannose-PTS family, which are phosphorylated on a histidine residue. IIB → P then phosphorylates the incoming sugar. The IIC domain is part of an integral membrane protein responsible for the translocation of the sugar molecule, and does not undergo phosphorylation. The EII complex of the mannose-PTS family, unlike the other EII complexes, has an additional domain called IID. This domain forms the permease in association with IIC (Saier & Reizer, 1992; Postma et al., 1993, 1996).

An analysis of several nucleotide sequences of EII genes allowed the PTS permeases to be classified into six families according to sequence similarity and domain organization (Postma et al., 1993; Lengeler et al., 1994). These families are: (a) glucose- and sucrose-PTS, (b) mannitol- and fructose-PTS, (c) lactose- and cellobiose-PTS, (d) mannose-PTS, (e) galactitol-PTS, and (f) galactitol-PTS. EIIIs of the mannose-PTS family are considered to be evolutionarily distinct from the other EII families (Reizer et al., 1996). The members of the mannose-PTS family that have been identified are EIIMan, EIINak and EIINak of Escherichia coli (Erni et al., 1987; Reizer et al., 1996), EIILev of Bacillus subtilis (Martin-Verstreete et al., 1990), EIILov of Klebsiella pneumoniae (Wehmeier & Lengeler, 1994), EIIMan/Glc of Vibrio furnissii (Bouma & Roseman, 1996) and EIIMan of Lactobacillus curvatus (Veyrat et al., 1996).

In Streptococcus salivarius, the mannose-PTS is constitutively produced (Vadeboncoeur, 1984), and is responsible for transporting mannose, glucose, fructose and 2-deoxyglucose (Vadeboncoeur, 1984). In addition to the permease moiety, this mannose-PTS was shown to contain two biochemically and antigenically related proteins, IIAMan and IIBMan, with molecular masses of 35.2 kDa and 38.9 kDa respectively (Bourassa et al., 1990). Both forms of IIAMan are present only in S. salivarius and Streptococcus vestibularis; 15 other species of streptococci, Lactococcus lactis and Lactobacillus casei possess only one protein that reacts with anti-IIAMan or IIBMan polyclonal antibodies. It has been shown that mutations affecting the expression of S. salivarius mannose-PTS components, more specifically IABAMan, have pleiotropic consequences on the expression of a wide variety of membrane and cytoplasmic proteins (Gauthier et al., 1990; Bourassa & Vadeboncoeur, 1992; Lapointe et al., 1993; Brochu et al., 1993), as well as on urease activity (Chen et al., 1998).

The biochemical characterization of S. salivarius IIABMan revealed that the IIA and IIB domains are phosphorylated on histidine residues, confirming that this enzyme belongs to the EII-mannose family (Pelletier et al., 1998). Further characterization also showed that IIABMan is required for the phosphorylation of mannose and 2-deoxyglucose by the mannose-PTS, while the IIABMan is unable to fulfill this function (Pelletier et al., 1998). The current work presents the isolation and characterization of the genes encoding the mannose-PTS of S. salivarius.

**METHODS**

**Bacterial strains, oligonucleotides, plasmids and growth conditions.** S. salivarius ATCC 25975 was cultivated in TYE medium (per litre: 10 g Tryptone, 5 g yeast extract, 2.5 g sodium chloride and 2.5 g disodium phosphate) supplemented with 0.5% (w/v) glucose. E. coli strains were cultivated aerobically in Luria–Bertani medium at 37 °C. Antibiotics were sterilized by filtration through 0.22 μm membranes (Amicon) and added aseptically to media to the following final concentrations: ampicillin, 50 μg ml⁻¹; tetracycline, 10 μg ml⁻¹; kanamycin, 10 μg ml⁻¹. Plasmid pUC18 (Yanisch-Perron et al., 1985) was used to construct genomic libraries, and pCRII (Invitrogen) was used for ampiclon cloning. The oligonucleotides used for PCR amplifications were: E3N29 (5′-ATGATGGATATATATGCTCGG-3′), which corresponded to the N-terminal sequence of IIABMan (Pelletier et al., 1998), E32205 (5′-TTGCGCGACATTGTCGCGTTTTTTC-3′), where Y = C or T, W = A or T, I = deoxynosine, which corresponded to the internal amino acid sequence of IIABMan, and the lacZ reverse universal primer RevlacZ (5′-GAGGGTAGAATACTTTACACAGG-3′).

The primers used to amplify the manL probe were 624 (5′-CAATTTAAAAACCGGAGGTATT-3′), which corresponded to the 219th basepair of manL, and 1011R (5′-GCCGGAACCCATATGATG-3′), which corresponded to the 605th basepair of manL. The manO probe was amplified using manO (5′-GTCCCGATACTAAATAAAG-3′), which corresponded to the 5′ end of manO, and manOR (5′-GAATTGAGTCTGGATCGCTGA-3′), which corresponded to the 328th basepair of manO. Amplicons were labelled using the random-primer [α-32P]dDNA labelling technique.

**DNA manipulations.** The genomic DNA of S. salivarius was isolated as previously described (Lortie et al., 1994). Unless otherwise mentioned, all DNA manipulations were performed using standard procedures (Ausubel et al., 1990). DNA sequencing was completed on both strands by the DNA sequencing service of Université Laval using appropriate synthesized oligonucleotides. Computer-assisted DNA and protein analyses were performed using the Genetics Computer Group Sequence Analysis software package, Version 9.1 (Devereux et al., 1984). At the time of writing, the relevant data from the genome sequencing project of Streptococcus mutans, Streptococcus pyogenes, Streptococcus pneumoniae and Enterococcus faecalis are available at http://www.ncbi.nlm.nih.gov/BLAST/unfinishedgenomed.html.

**RNA manipulations.** Total RNA was extracted from S. salivarius as previously described (Gagnon et al., 1995). Briefly, S. salivarius cells were harvested at mid-exponential phase by centrifugation and resuspended in 4 ml ice-cold 10 mM Tris (pH 8.0) containing 1 mM EDTA. Resuspended cells were mixed with a solution containing 3.5 g glass beads (150–212 μm diameter, Sigma), 3 ml phenol/chloroform/isooamyl alcohol (100:24:1, by vol.) and 0.25 ml 10% (w/v)
SDS, and vortexed for 3 min at 4 °C. Lysed cells were centrifuged and the supernatant was extracted five times with phenol/chloroform/isoamyl alcohol (100:24:1, by vol.). Ribonucleic acids were selectively precipitated by adding 0.1 vol. 10 M LiCl and 2 vols 95% ethanol. The RNA pellet was washed with 70% ethanol, air-dried, resuspended in 500 μl diethylpyrocarbonate-treated water, and stored in aliquots at −80 °C.

Denaturing formaldehyde agarose gels (1 %, w/v), buffers and samples were prepared according to Ausubel et al. (1990). RNA from S. salivarius ATCC 25975 (10 μg) was transferred to positively charged nylon membranes (Roche) using a Possiblot pressure blotter according to the manufacturer’s instructions (Stratagene). Total RNA was fixed by UV cross-linking. RNA molecular mass markers (Gibco-BRL) were used to determine the size of the transcripts. Primer extension analysis was performed as described by Ausubel et al. (1990) using primer 424R (5'-AATAATACCGGATACCCATTCGT-GTT-3') labelled with T4 polynucleotide kinase and [γ-32P]ATP. The radiolabelled oligonucleotide (10 ng) was hybridized with 20 μg S. salivarius total RNA, and the extension was performed using 200 U murine leukaemia virus (MLV) reverse transcriptase (Gibco-BRL) for 1 h at 42 °C. The extended product was denatured and analysed by electrophoresis on a 9% polyacrylamide gel containing 7 M urea.

**Western blot analysis.** Crude extracts of E. coli XL-1 Blue bearing pML17D or pML18D were prepared by sonication as previously described (Gagnon et al., 1995). Membrane-free cell extracts of S. salivarius were submitted to SDS-PAGE, electrotransferred to nitrocellulose membranes, and probed using rabbit anti-IIAB<sup>Man</sup><sub>L</sub> polyclonal antibodies as previously described (Pelletier et al., 1995).

**RESULTS AND DISCUSSION**

**Cloning of the gene encoding IIAB<sup>Man</sup><sub>L</sub>**

IIAB<sup>Man</sup><sub>L</sub> was extracted from two-dimensional gels and the sequence of the N-terminal amino acid residues was determined by Edman degradation. The resulting sequence was MIGIIIASHGKFAEAG (Pelletier et al., 1998). The purified IIAB<sup>Man</sup><sub>L</sub> was subsequently submitted to CnBr proteolysis, and the amino acid sequences of internal fragments were determined. The sequence of one internal peptide corresponded to ANATAEQVAANII. Two degenerate oligonucleotides, E3N29 and E3I2205 (see Methods), corresponding respectively to the N-terminal and the internal amino acid sequences, were designed. PCR using these oligonucleotides generated a 364 bp amplicon from the S. salivarius chromosome. Enriched genomic libraries of S. salivarius cloned into pUC18 failed to produce clones that hybridized with the 364 bp probe, suggesting that the chromosomal fragment bearing the IIAB<sup>Man</sup><sub>L</sub>-encoding gene and/or the IIAB<sup>Man</sup><sub>L</sub> protein was toxic for E. coli cells.

To circumvent this toxicity problem, PCR using the E3N29 primer and RevlacZ universal primer was performed on a ligation mixture of a 4–7 kb Tsp509I partial genomic library of S. salivarius cloned into pUC18. The amplicons were ligated to PCR-II (Invitrogen) and transformed into E. coli XL-1 Blue. Screening using the 364 bp amplicon as a probe allowed the identification of one clone (pML17D) carrying a 1.2 kb amplicon. Analysis of the nucleotide sequence revealed the presence of 990 bp ORF encoding a 330 aa protein that shared a high degree of identity with IIAB<sup>Man</sup><sub>H</sub> and IIAB<sup>Man</sup><sub>L</sub> proteins of the mannose-PTS family. The predicted molecular mass of this putative protein was 35.5 kDa, which was almost identical to the molecular mass of IIAB<sup>Man</sup><sub>L</sub> (35.2 kDa) determined by SDS-PAGE (Bourassa et al., 1990). Moreover, E. coli cells bearing pML17D produced a 35 kDa protein, which was recognized by anti-IIAB<sup>Man</sup><sub>L</sub> antibodies. This protein was not found in extracts from E. coli cells bearing pML18D, which carried the same DNA fragment inserted in the opposite orientation (Fig. 1). This result proved that IIAB<sup>Man</sup><sub>L</sub> is not toxic for E. coli under these growth conditions. The 35 kDa protein from pML17D was phosphorylated in the presence of PEP, EI and HPr proteins of S. salivarius, and allowed the phosphorylation of 2-deoxyglucose in the presence of membrane fractions from a S. salivarius mutant (G77) lacking IIAB<sup>Man</sup><sub>L</sub> and IIAB<sup>Man</sup><sub>H</sub> (Pelletier et al., 1998). These results confirmed that the product of the gene (manL) is the IIAB<sup>Man</sup><sub>L</sub> of S. salivarius.

**manL is part of a tetracistronic operon**

A Northern blot analysis of total RNA isolated from S. salivarius using a manL coding portion as a DNA probe allowed the detection of a single 3.6 kb mRNA transcript...
L.-A. LORTIE and OTHERS

ORFs, which we named manL, manM, manN and manO. Two terminator-like structures flanked the man operon: one was located 155 bp upstream from the Pman promoter, and the second was 128 bp downstream from manO. The 5’ end of an ORF, highly similar to the seryl-tRNA synthetase gene (serS) of Staphylococcus aureus (40% amino acid identity) and transcribed in the same orientation as the man operon, was located downstream from manO. An analysis of sequences currently available in the databases revealed that the same gene organization is found in S. pyogenes, i.e. manL, manM, manN, manO, serS, while in S. mutans, discontinuity of the sequences impeded complete interpretation of the 5’ end of the operon. However, analysis of the 3’ end of the man operon revealed that a 3–9 kb fragment bearing the comA gene is present between manN and manO, and that a 1–8 kb fragment bearing no identified genes is present between manO and serS. These data suggest that this region of the genome of S. mutans was originally identical to that of S. salivarius and S. pyogenes, but was subsequently subjected to chromosomal rearrangement. However, we cannot exclude the possibility that the comA gene was present in the chromosome of S. salivarius and S. pyogenes and was subsequently excised.

In the S. pneumoniae genome, the gene encoding the ManO homologue is not contiguous with PTS-related genes, but is downstream from the thrA gene encoding the aspartokinase. As previously described, the manO gene of S. mutans was separated from the man operon by the insertion of a 3–9 kb fragment bearing the comA gene. The chromosomal environment is unknown for the Lc. lactis gene. Northern blot analysis using the manO coding region as a DNA probe revealed the presence of two mRNAs, a long 3–6 kb transcript originating from Pman and encompassing the entire man operon (Fig. 3). The man operon was composed of four ORFs, which we named manL, manM, manN and manO. Two terminator-like structures flanked the man operon: one was located 155 bp upstream from the Pman promoter, and the second was 128 bp downstream from manO. The 5’ end of an ORF, highly similar to the seryl-tRNA synthetase gene (serS) of Staphylococcus aureus (40% amino acid identity) and transcribed in the same orientation as the man operon, was located downstream from manO. An analysis of sequences currently available in the databases revealed that the same gene organization is found in S. pyogenes, i.e. manL, manM, manN, manO, serS, while in S. mutans, discontinuity of the sequences impeded complete interpretation of the 5’ end of the operon. However, analysis of the 3’ end of the man operon revealed that a 3–9 kb fragment bearing the comA gene is present between manN and manO, and that a 1–8 kb fragment bearing no identified genes is present between manO and serS. These data suggest that this region of the genome of S. mutans was originally identical to that of S. salivarius and S. pyogenes, but was subsequently subjected to chromosomal rearrangement. However, we cannot exclude the possibility that the comA gene was present in the chromosome of S. salivarius and S. pyogenes and was subsequently excised.

In the S. pneumoniae genome, the gene encoding the ManO homologue is not contiguous with PTS-related genes, but is downstream from the thrA gene encoding the aspartokinase. As previously described, the manO gene of S. mutans was separated from the man operon by the insertion of a 3–9 kb fragment bearing the comA gene. The chromosomal environment is unknown for the Lc. lactis gene. Northern blot analysis using the manO coding region as a DNA probe revealed the presence of two mRNAs, a long 3–6 kb transcript originating from Pman and encompassing the entire man operon, and a small 0.7 kb mRNA (Fig. 2b). This result confirms that manO is transcribed not only from Pman but also from another promoter in the manN–manO intergenic region. The putative PmanO had a consensus —10 box, while its putative —35 box had two mismatches with the consensus sequence, and was

Fig. 2. (a) Northern blot of total S. salivarius RNA using manL as a DNA probe. The arrow indicates the 3.6 kb man operon transcripts. (b) Northern blot of total S. salivarius RNA using manO as a DNA probe. The arrows indicate the 3.6 kb and 0.7 kb transcripts.

Fig. 3. Schematic representation of the S. salivarius man operon. Promoters (P) and putative terminators (T) are indicated. The manL- and manO-specific DNA probes used in this study are indicated by black boxes above the corresponding ORFs.
The asterisk indicates the 5' terminus of the coding region of the man operon. Other authors have previously reported that DNA fragments from streptococci are toxic for *E. coli* due to the presence of strong promoters or promoter-like sequences (Stassi & Lacks, 1982; Chen & Morrison, 1987; Martin et al., 1989; Dillard & Yother, 1991).

The coding region for manL is located 117 bp downstream from the transcriptional startpoint of the man operon. This portion of the transcript did not have an ORF preceded by an RBS. This 5'-untranslated region (5'-UTR) has several stretches of adenines and thymines, and is richer in A+T (70 mol%) than the rest of the chromosome (60 mol%). In *silico* studies revealed that 5'-UTR can form several alternative secondary structures with few unpaired nucleotides and energy levels as high as 29.5 kcal mol⁻¹ (123 kJ mol⁻¹) (results not shown). 5'-UTR sequences were detected in several mRNAs from different bacteria like *E. coli*, *Rhodobacter capsulatus* and *B. subtilis* (for reviews see Nierlich & Murakawa, 1996; Kushner, 1996; Coburn & Mackie, 1999), and mRNAs from cold-shock genes from *L. lactis* (Wouters et al., 1998) and *E. coli* (Jiang et al., 1996; Fang et al., 1998). These structures have been shown to help stabilise the transcript (Coburn & Mackie, 1999) and to be important in the regulation of gene expression (Fang et al., 1998). The function of this untranslated mRNA region is currently under investigation.

**Characterization of the genes composing the man operon**

**manL**. The percentages of identity of the inferred amino acid sequence of IIAB\textsubscript{L} man with the corresponding domains of homologous proteins are presented in Table 1. It is interesting to note that the IIB domains share a slightly higher degree of conservation than the IIA domains, possibly reflecting structural constraints imposed for recognition of the IIC and IID mannose permeases. Analyses of sequences available in the databases revealed that in addition to *S. salivarius* and *E. coli*, genes encoding IIAB\textsubscript{L} man polypeptides are also present in the genomes of *S. mutans*, *S. pyogenes*, *S. pneumoniae* and *Ent. faecalis*. The IIA and the IIB domains of other mannose-PTSs are on separate polypeptides. A comparison of the amino acid sequences surrounding the IIA phosphorylation sites revealed a high level of identity (Fig. 5a). These levels of identity are particularly high among IIAAs from streptococci and enterococci. The amino acid sequences of the IIA domains from *E. coli* and *S. salivarius* only shared 35% identity (Table 1, Fig. 5a). Nevertheless, their structures have numerous points in common. Among the residues determined to be in the vicinity of the IIA phosphorylated histidine (H₁₀) of the *E. coli* IIAB\textsubscript{L} man by crystallographic studies (M₂₃, L₄₄, D₆₇, G₇₁, S₇₂ and P₇₃) (Nunn et al., 1996), several were conserved in the *S. salivarius* IIAB\textsubscript{L} man (M₂₃, L₄₄, D₆₇, G₇₁, T₇₂ and P₇₃). The hydroxyl group of S₇₂, which was shown to be essential for PTS activity (Nunn et al., 1996), is replaced in the S.

Fig. 4. Mapping of the 5’-terminus of the man operon transcript. Primer extension experiments were performed using 20 µg *S. salivarius* RNA. The end-labelled primer (424R; lane P) was complementary to the beginning of the coding region of the manL gene. Lanes A, C, G, T correspond to the dideoxy sequencing reactions performed using the same oligonucleotide primer. The asterisk indicates the 5’ terminus of the man transcript.
Table 1. Level of identity of S. salivarius IIAB<sup>Man<sub>sal</sub></sup> with the IIA<sup>Man</sup> and IIB<sup>Man</sup> domains from various bacteria

<table>
<thead>
<tr>
<th>Organism</th>
<th>IIA domain</th>
<th>IIB domain</th>
<th>Source or accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Residues</td>
<td>Identity (%)</td>
<td>Residues</td>
</tr>
<tr>
<td>S. salivarius</td>
<td>1–150</td>
<td>100</td>
<td>151–330</td>
</tr>
<tr>
<td>S. pneumoniae</td>
<td>1–145†</td>
<td>83·1</td>
<td>146–322</td>
</tr>
<tr>
<td>S. pyogenes</td>
<td>1–150</td>
<td>86·0</td>
<td>151–330</td>
</tr>
<tr>
<td>S. mutans</td>
<td>1–130‡</td>
<td>84·0</td>
<td>NA</td>
</tr>
<tr>
<td>Ent. faecalis</td>
<td>1–150</td>
<td>56·1</td>
<td>151–342</td>
</tr>
<tr>
<td>Lb. curvatus</td>
<td>1–146</td>
<td>28·0</td>
<td>1–162</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>1–146</td>
<td>25·7</td>
<td>1–163</td>
</tr>
<tr>
<td>E. coli (man)</td>
<td>1–150</td>
<td>35·0</td>
<td>151–322</td>
</tr>
<tr>
<td>E. coli (aga)</td>
<td>NA</td>
<td>–</td>
<td>1–158</td>
</tr>
<tr>
<td>E. coli (aga')</td>
<td>NA</td>
<td>–</td>
<td>1–157</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td>1–135</td>
<td>27·6</td>
<td>1–150</td>
</tr>
<tr>
<td>V. furnissii</td>
<td>1–157</td>
<td>30·5</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA. Not applicable.

† Incomplete sequence at 5' end of the corresponding gene.
‡ Incomplete sequence at 3' end of the corresponding gene.

S. salivarius homologue by another OH-bearing residue, T<sub>24</sub>. In addition, in the S. salivarius IIAB<sup>Man</sup>sal, the hydrophobic L<sub>24</sub> residue is conservatively replaced by I<sub>24</sub>. Consequently, the molecular environment of the IIA phosphorylation site of the S. salivarius IIAB<sup>Man</sup>sal should be highly similar to that of the E. coli of IIAB<sup>Man</sup>. The amino acid sequences in the vicinity of the IIB phosphorylation sites also showed high levels of identity, especially for proteins from streptococci and enterococci (Table 1, Fig. 5a). It is important to note that the amino acid sequences of the two inferred phosphorylation sites of IIAB<sup>Man</sup>sal corresponded to those determined from the biochemical approach (Pelletier et al., 1998). The five arginine residues reported to be involved in the mannose-PTS activity of the E. coli IIIB<sup>Man</sup> (<sup>R</sup><sub>658</sub>, <sup>R</sup><sub>1372</sub>, <sup>R</sup><sub>190</sub>, <sup>R</sup><sub>204</sub> and <sup>R</sup><sub>331</sub>) (Gutknecht et al., 1998) are strictly conserved, with the exception of IIB<sup>Ag</sup>a and IIB<sup>Ag</sup>a. Other workers have proposed a signature sequence for the IIB domain of the mannose-PTS family (Reizer et al., 1996). This signature ([LIVM]DXR[LIVM]F_HGQ-[LIVM]X,W) corresponds perfectly to the IIB phosphorylation domain of IIAB<sup>Man</sup>sal. Taken together, these results confirm that IIAB<sup>Man</sup>sal of S. salivarius is unequivocally a member of the mannose-PTS family.

<sup>manM</sup>. The <sup>manM</sup> gene was located 67 bp downstream from <sup>manL</sup> and coded for a hydrophobic 271 aa polypeptide (27·2 kDa). An analysis of database sequences revealed that <sup>manM</sup> shared identity with the mannose-PTS IIC domains of S. pyogenes (79%), S. pneumoniae (77%), Ent. faecalis (71%), S. mutans (68%), E. coli (46%), Lb. curvatus (44%) and K. pneumoniae (41%), and with B. subtilis levF (43%), E. coli agaC (27%) and E. coli agaC (23%). A putative RBS (AAAGGA) is present 9 bp upstream from the <sup>manM</sup> start codon. Sequence alignments of homologous IIC<sup>Man</sup> proteins from different bacteria revealed that: (i) IIC<sup>Man</sup> of streptococci and enterococci share high levels of identity (71–79%) and (ii) the residues that were previously identified by Reizer et al. (1996) as being strictly conserved among the IIC<sup>Man</sup> enzymes of the man, aga and aga' systems of E. coli and the lev system of B. subtilis (G<sub>202</sub>, P<sub>285</sub>, G<sub>411</sub>, G<sub>153</sub>, G<sub>501</sub>, G<sub>552</sub>, P<sub>773</sub>, D<sub>120</sub>, P<sub>176</sub>, G<sub>182</sub>, G<sub>187</sub>, V<sub>193</sub>, G<sub>194</sub>, M<sub>202</sub>, P<sub>209</sub>, G<sub>214</sub>, F<sub>257</sub>, L<sub>288</sub>, G<sub>223</sub>, A<sub>257</sub>) were also conserved in the IIC<sup>Man</sup> sal or pneumoniae, S. mutans, S. pyogenes, S. pneumoniae, Ent. faecalis and Lb. curvatus (data not shown). This strengthens the hypothesis that these residues are important for the structure and/or function of the enzyme. However, other residues reported by the same authors to be conserved in IIC<sup>Man</sup> (L<sub>441</sub>, Q<sub>453</sub>, R<sub>457</sub>, L<sub>441</sub>, D<sub>460</sub>, T<sub>490</sub>, E<sub>558</sub>, G<sub>633</sub>, Q<sub>695</sub>, T<sub>805</sub>, G<sub>188</sub> and Y<sub>240</sub>) were not maintained among streptococci, enterococcal and lactobacillus homologues, suggesting that these residues are subject to more important evolutionary divergence than previously believed. None of the 12 variations occurred in the central cytoplasmic loop of IIC<sup>Man</sup> (Huber & Erni, 1996; Reizer et al., 1996). Among the 13 nonconserved residues, four (L<sub>441</sub>, D<sub>460</sub>, T<sub>490</sub>, E<sub>558</sub>) were part of the IIC<sup>Man</sup> family signature proposed by Reizer et al. (1996) (Fig. 5b). Taking into account the sequences of Gram-positive bacteria genomes currently available, we propose a modified IIC<sup>Man</sup> signature corresponding to GX,G[DNH][X,G][LIVM],[X,G][-STL][-LT][-EQ]. Reizer et al. (1996) reported that E<sub>558</sub> is a candidate for a crucial role in enzyme activity, and could correspond to the glutamyl residue involved in sugar binding and the catalytic function of IIC proteins (Jacobson & Saraceni-
The *man* operon of *Streptococcus salivarius*

<table>
<thead>
<tr>
<th>IIA</th>
<th>IIB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cons</td>
<td>H G</td>
</tr>
<tr>
<td>Ssa man</td>
<td>1 MGIGIIAHSKHGFARHGSQGSGMFG</td>
</tr>
<tr>
<td>Smu man</td>
<td>1 MAIGIVSHAEFGAASIQSGGSGMFG</td>
</tr>
<tr>
<td>Spy man</td>
<td>1 MGIGIIAHSKHGFARHGSQGSGMFG</td>
</tr>
<tr>
<td>Spa man</td>
<td>1 MGIGIIAHSKHGFARHGSQGSGMFG</td>
</tr>
<tr>
<td>Eco man</td>
<td>1 MTIAIVGTGHGWAQLELKTAEMLG</td>
</tr>
<tr>
<td>Kpn sor</td>
<td>1 NVHAIFCAHQLGALMDSVCWYG</td>
</tr>
<tr>
<td>Eco aga'</td>
<td>1 NVHAIFCAHQLGALMDSVCWYG</td>
</tr>
</tbody>
</table>

**Fig. 5.** (a) Alignment of the IIA and IIB phosphorylation sites of the mannose-PTS family. The phosphorylated histidine residues are in bold and the signature sequence is boxed. (b) Alignment of the IIC domain portions located in the vicinity of the signature sequence. The glutamate residue previously reported to be essential is in bold. (c) Alignment of the IID domain portions located in the vicinity of the signature sequence. The strictly conserved glutamate residue (E29) is in bold and the KLTEG motif is underlined. Strictly conserved residues are indicated in the consensus line (Cons) by capital letters. Residues conserved among the Gram-positive bacteria are indicated by ‗>‘. The signature region is boxed. References for the sequences from the following organisms are presented in Table 1: Ssa, *Streptococcus salivarius*; Smu, *Streptococcus mutans*; Spy, *Streptococcus pyogenes*; Spn, *Streptococcus pneumoniae*; Efa, *Enterococcus faecalis*; Lcu, *Lactobacillus curvatus*; Bsu, *Bacillus subtilis*; Eco, *Escherichia coli*; Kpn, *Klebsiella pneumoniae*; Vfu, *Vibrio furnissii*.

Richards, 1993). This glutamate is substituted by a glutamine in the IIC\textsuperscript{Man}\textsuperscript{+} of the six Gram-positive species, suggesting that a negative charge at this position does not appear to be essential for IIC activity. Interestingly, the IIC\textsuperscript{Man} of all the Gram-positive bacteria possess strictly conserved glutamate (E\textsubscript{142}) and aspartate (D\textsubscript{130}) residues that might compensate for the absence of E\textsubscript{9} and maintain PTS activity. The Cys\textsubscript{37} residue proposed to be involved in the interactions of IIC\textsuperscript{Man} and IID\textsuperscript{Man} (Rhiel et al., 1994) is conserved in IIC\textsuperscript{Man} proteins. However, despite its ubiquitous presence in IIC\textsuperscript{Man} proteins, mutagenesis of this residue in the *E. coli* IIC\textsuperscript{Man} demonstrated that it is not essential for catalytic activity (Rhiel et al., 1994).

\textbf{manN.} The manN gene is located 14 bp downstream from the manM stop codon. It encodes a putative 303 aa protein (33-4 kDa), which shared strong identity with the mannose-PTS IID domains of *S. pyogenes* (80%), *S. pneumoniae* (74%), *S. mutans* (74%), *Ent. faecalis* (68%), *E. coli* (55%), *K. pneumoniae* (49%), and with *B. subtilis* levG (48%) and *E. coli* agaD (37%). The gene

---

**Table 1:**

<table>
<thead>
<tr>
<th>Organism</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ssa</td>
<td>Streptococcus salivarius</td>
</tr>
<tr>
<td>Smu</td>
<td>Streptococcus mutans</td>
</tr>
<tr>
<td>Spy</td>
<td>Streptococcus pyogenes</td>
</tr>
<tr>
<td>Spn</td>
<td>Streptococcus pneumoniae</td>
</tr>
<tr>
<td>Efa</td>
<td>Enterococcus faecalis</td>
</tr>
<tr>
<td>Lcu</td>
<td>Lactobacillus curvatus</td>
</tr>
<tr>
<td>Bsu</td>
<td>Bacillus subtilis</td>
</tr>
<tr>
<td>Eco</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>Kpn</td>
<td>Klebsiella pneumoniae</td>
</tr>
<tr>
<td>Vfu</td>
<td>Vibrio furnissii</td>
</tr>
</tbody>
</table>
is preceded by a putative RBS (GGGGTG) 5 bp upstream from the ATG start codon. The KLTEG motif reported to be involved in permease–sugar interactions (Jacobson & Saraceni-Richards, 1993; Lengeler et al., 1994; Wehmeier & Lengeler, 1994) was not conserved in streptococcal IID enzymes (Fig. 5c), as in IID\textsuperscript{Man} of *E. coli* (Reizer et al., 1996). The motif was replaced by a [DN]ITKG motif. A careful comparison of the motifs showed that the L residue was replaced by an I residue, a modification that could be regarded as conservative. Furthermore, with the exception of the *S. mutans* protein, the positively and negatively charged residues were inverted (K$^{+}$LTE$^{-}$G for D[TG]$^{+}$K$^{-}$), leaving open the possibility that these residues are equally involved in the permease–sugar interaction. A highly conserved region (residues 24–48) on the cytoplasmic side of the permease (Huber & Erni, 1996) included a strictly conserved glutamate residue (E\textsubscript{29}), which might be involved in the enzyme–substrate interaction. Interestingly, the IID\textsuperscript{Man} of several Gram-positive bacteria contains an additional segment of about 30 aa at the cytoplasmic-membrane junction. The previously reported IID\textsuperscript{Man} family signature ([K[LIVM][G]LIVM], GP[LIVM]AG[LIVM]GD[P][LIVMF]W) (Reizer et al., 1996) was strictly conserved among the IID\textsuperscript{Man} domains presented in Fig. 5(c). Thus, *manM* and *manN* encode the permease portions of the mannosone-PTS system of *S. salivarius*.

**manO.** The fourth and last gene of the *man* operon, *manO* is located 78 bp downstream from *manN*, and encodes a small basic (estimated pl 10.7) 124 aa protein (13.7 kDa). The basic pl suggested that it might interact with nucleic acids. However, in *silico* structure analyses of the ManO protein failed to assign the classical helix–turn–helix motif to a portion of the polypeptide. A BLAST search of databases revealed high levels of identity with putative proteins from *S. pneumoniae* (79%), *S. pyogenes* (65%), *S. mutans* (64%), *Ent. faecalis* (32%) and *L. lactis* (33%). No function has been assigned to these proteins, but the genes are also located downstream from the IID\textsuperscript{Man} genes in *S. pyogenes* and *Ent. faecalis*.

**Conclusions**

We have reported the first isolation and characterization of a complete *man* operon for a Gram-positive eubacterium. This operon was comprised of four genes encoding IID\textsuperscript{Man}, a IID\textsuperscript{Man}, a IID\textsuperscript{Man} and a protein of unknown function. Identification of the precise functions of ManO and the putative regulatory roles of IID\textsuperscript{Man} is currently under investigation using insertional mutagenesis.

**ACKNOWLEDGEMENTS**

This research was supported by Medical Research Council of Canada (MRC) operating grants MT-6979 to C.V. and MT-11276 to M.F., and the Fonds pour la Formation de Chercheurs et l’Aide à la Recherche from the Province of Québec. M.F. is a scholar of the Fonds de la Recherche en Santé du Québec. We thank France Dumas for performing the CnBr proteolysis digestions and sequencing the peptides (Biotechnology Research Institute, National Research Council of Canada).

**REFERENCES**


684


