
Doctorate of Philosophy
The University of Sydney

The Glucosyltransferases of
Streptococcus salivarius

Christine L. Simpson
1994

DENTISTRY LIBRARY
UNIVERSITY OF SYDNEY

CONTENTS

PREFACE	(viii)
STATEMENT OF AUTHORSHIP	(ix)
ABSTRACT	(x)
ABBREVIATIONS	(xi)
1 GLUCOSYLTRANSFERASES OF ORAL STREPTOCOCCI	1
1.1 Structure of Streptococcal Glucans	1
1.2 Biosynthesis of Glucans	2
1.2.1 Reaction 1: Glucosyl Transfer to an Acceptor	3
1.2.2 Reaction 2: Autopolymerization	6
1.2.3 Reaction 3: Hydrolysis of sucrose	7
1.2.4 Reaction 4: Isotope Transfer	7
1.3 Isozymes of Glucosyltransferase	7
1.4 Cloning of Glucosyltransferase Activities	8
1.5 The Domain Structure of the Glucosyltransferases	10
1.5.1 The Signal Peptide	11
1.5.2 The 'Variable' Domain	12
1.5.3 The 'Catalytic' Domain	12
1.5.4 The Glucan-Binding Domain	12
1.6 The Catalytic Site of the Glucosyltransferases	15
1.7 Glucan Binding Proteins	17
1.8 Regulation of Glucosyltransferases	20
1.8.1 Effect of Carbon Source and Growth Rate	20
1.8.2 Effect of Tween 80	21
1.8.3 Effect of [Na ⁺]/[K ⁺]	22
1.8.4 Effect of Other Compounds	22
1.8.5 Regulatory Element of <i>S. gordonii</i> Challis	24

1.9	Caries and Colonization	24
1.9.1	Role of Glucosyltransferases in Adhesion	25
1.9.3	Dextran-Mediated Interbacterial Aggregation	28
1.9.4	GTF Phase Variation of <i>S. gordonii</i>	28
1.9.5	Sucrose-Independent Adhesion	29
1.10	Effect of Antiplateau Agents	30
1.10.1	Glucanases as Antiplateau Agents	30
1.10.2	Inhibitors of Glucosyltransferase Activity	31
1.11	A Role for Glucan in Endocarditis	32
1.12	Glucosyltransferases as Potential Immunogens	33
2	ROUTINE METHODS	35
2.1	Reagents, Chemicals and Enzymes	35
2.2	Bacterial Strains, Phages and Phagemids	37
2.3	Media and Growth Conditions	37
2.4	DNA Manipulations	40
2.4.1	Small Scale Purification and Isolation of Bacteriophage- λ DNA	40
2.4.2	Large Scale Isolation and Purification of Bacteriophage- λ DNA	40
2.4.3	Rapid Isolation and Purification of Phagemid DNA . .	41
2.4.4	Large Scale Isolation and Purification of Phagemid DNA	41
2.4.5	Preparation of Single Stranded Phagemid DNA	42
2.4.6	Spectrophotometric Determination of Nucleic Acid Concentration	43
2.4.7	Precipitation of DNA by Ethanol	43
2.4.8	Digestion of DNA with Restriction Endonucleases	43

2.4.9	Gel Electrophoresis of DNA on Agarose Gels	44
2.4.10	Dephosphorylation	44
2.4.11	Recovery of DNA from Agarose Gels	44
2.4.12	Ligation	45
2.4.13	Preparation of Cells for Electroporation	45
2.4.14	Transformation by Electroporation	45
2.4.15	Transfer of DNA to Nylon Membranes for Southern Hybridizations	46
2.4.16	Capillary Blotting of Agarose Gels	46
2.4.17	Hybridizations	46
2.4.18	DNA Sequencing	47
2.5	Detection of Glucosyltransferase	
	Activity	47
2.5.1	Detection of Sucrase Activity	47
2.5.2	Detection of Glucan Synthesis	48
2.5.3	SDS-PAGE Gel Electrophoresis for the Detection of Glucosyltransferase Activity	48
2.5.4	Detection of Glucosyltransferase Activity on SDS- PAGE Gels	49
3	ISOLATION OF RNA AND TRANSCRIPTIONAL ANALYSIS OF THE <i>gtfJK</i> LOCUS	50
3.1	Introduction	50
3.2	Materials and Methods	52
3.2.1	RNA Isolation from <i>S. salivarius</i> ATCC 25975	52
3.2.2	RNA Gel Electrophoresis	53
3.2.3	Northern Hybridizations	53
3.2.4	RNase Protection Assays	54
3.3	Results	55
3.3.1	RNA Isolation	55

3.3.2	Transcriptional Analysis of the <i>gtfJK</i> Locus	57
3.3.3	Detection of a Putative <i>gtfJ</i> Transcriptional Terminator	57
3.4	Discussion	59
4	ANALYSIS OF THE REGION UPSTREAM OF THE <i>gtfJK</i> LOCUS	61
4.1	Introduction	61
4.2	Methods and Materials	61
4.2.1	Blotting of Bacteriophage- λ Gene Bank	61
4.2.2	Plaque and Southern Hybridizations	63
4.2.3	Production of Bacteriophage- λ Lysates	64
4.2.4	DNA Sequence Analysis	64
4.3	Results and Discussion	64
4.3.1	Sequencing Upstream of the <i>gtfJK</i> Locus	64
4.3.2	Detection of <i>gtfX</i> Positive Plaques	66
4.3.3	Restriction Mapping of λ B-5.2	66
4.3.4	Evidence for a Mixed Insert in λ B-5.2	68
4.3.5	Sequence Similarity Between Regions of the Insert in λ B-5.2 and that in λ A-8	68
4.3.6	Does the <i>gtfX</i> gene exist?	71
5	CLONING OF TWO GENES FROM <i>STREPTOCOCCUS SALIVARIUS</i> ENCODING PRIMER-INDEPENDENT GLUCOSYLTRANSFERASE ACTIVITIES	72
5.1	Introduction	72
5.2	Methods and Materials	73
5.2.1	Screening for Clones Expressing Glucosyltransferase Activity	73
5.2.2	Production of Lysates	73

5.2.4	Digestion of Glucan with Dextranase	74
5.2.5	Initial Characterization of the Carbohydrate Products Synthesized by the λ D-10 Lysate	74
5.2.6	Southern and Northern Hybridizations	75
5.2.7	SDS-PAGE and Isoelectric Focusing	76
5.3	Results	77
5.3.1	Detection of Polymer-Forming Bacteriophage- λ on Sucrose Agar Plates	77
5.3.2	Preliminary Analysis of Glucans Formed by the Various GTF Activities of <i>S. salivarius</i>	80
5.3.3	Endonuclease Restriction Maps of the Chromosomal DNA in λ D-10 and λ C-13	82
5.3.4	Subcloning of the <i>gtf</i> Genes	82
5.3.5	Southern and Northern Hybridization Analysis	82
5.3.6	SDS-PAGE and Isoelectric Focusing Analysis	84
5.4	Discussion	88
6	SEQUENCE ANALYSIS OF THE <i>gtfL</i> AND <i>gtfM</i> GENES	95
6.1	Introduction	95
6.2	Methods	97
6.2.1	DNA Sequence Determination	97
6.2.2	Multiple Alignments and Phylogenetic Tree Construction	97
6.3	Results and Discussion	99
6.3.1	Analysis of <i>gtfL</i>	99
6.3.1.1	Nucleotide Sequence	99
6.3.1.2	Deduced Amino Acid Sequence	99
6.3.1.3	Repeat Units of GtfL	107
6.3.2	Analysis of <i>gtfM</i>	114
6.3.2.1	Nucleotide Sequence	114
6.3.2.2	Deduced Amino Acid Sequence	114

6.3.2.3	Repeat Units of GtfM	122
6.3.3.	Multiple Alignment and Phylogenetic Comparison with other Gtf Sequences	126
7	¹³ C-NMR ANALYSIS OF THE GLUCANS PRODUCED BY THE GLUCOSYLTRANSFERASES OF <i>S. SALIVARIUS</i>	144
7.1	Introduction	144
7.2	Methods and Materials	145
7.2.1	Partial Purification of Recombinant Glucosyltransferases	145
7.2.3	Preparation of Glucan	145
7.2.4	¹³ C-NMR analysis of Glucans	146
7.3	Results and Discussion	146
7.3.1	Partial Purification of Recombinant Glucosyltransferase Activities	146
7.3.2	Analysis of the Glucans Produced by <i>S. salivarius</i> Glucosyltransferases	146
7.3.2.1	¹³ C-NMR Analysis of Dextran T-10	151
7.3.2.2	¹³ C-NMR analysis of the Glucan Produced by GtfJ	151
7.3.2.3	¹³ C-NMR Analysis of the Glucan Produced by GtfK	158
7.3.2.4	¹³ C-NMR analysis of the Glucan Produced by GtfL	158
7.3.2.5	¹³ C-NMR Analysis of the Glucan Produced by GtfM	165
8	KINETIC PROPERTIES OF GLUCOSYLTRANSFERASES EXPRESSED IN <i>ESCHERICHIA COLI</i>	168
8.1	Introduction	168
8.2	Methods and Materials	169
8.2.1	Preparation of Samples	169
8.2.2	Determination of the Amount of Glucose Incorporated into Glucan	169

8.2.3 Production of Glucan with Time 169

8.2.4 Effect of Dextran Concentration on
Glucosyltransferase Activity 170

8.2.5 Effect of Sucrose Concentration on
Glucosyltransferase Activity 170

8.2.6 Effect of pH on Glucosyltransferase Activity 170

8.2.6 Glucan Products as Acceptors for Primer-
Dependent Activities 171

8.3 Results and Discussion 171

8.3.1 Preliminary Data 171

8.3.2 Production of Glucan with Time 174

8.3.3 Effect of Dextran Concentration on Glucan
Production 178

8.3.4 Effect of Sucrose Concentration on Glucan
Production 184

8.3.5 Effect of pH on Glucan Production 191

8.3.6 Acceptor Efficiency of Glucans Produced by Primer-
Independent Activities 191

9 DISCUSSION 204

APPENDIX 1 PREPARATION OF SEMI-DEFINED MEDIA (SDM) 216

BIBLIOGRAPHY 217

PREFACE

I would like to express my thanks to my supervisor, Dr Nick Jacques, and to Dr Phil Giffard for their support and guidance during the course of this work. I would also like to thank Professor Ken Knox, Director of the Institute of Dental Research, for making available the facilities of the Institute during this course of study.

I would also like to thank my husband, David, and my family for their encouragement and assistance.

This work was funded by an NHMRC Program Grant awarded to Dr Nick Jacques and Professor Ken Knox at the Institute of Dental Research.

Some of the work described in this thesis has already been reported:

Simpson, CL, Giffard PM, Jacques NA (1993). A method for the isolation of RNA from *Streptococcus salivarius* and its application to the transcriptional analysis of the *gtfJK* locus. *FEMS Microbiol. Letts.* **108**: 93-98.

Simpson CL, Giffard PM, Jacques NA (1991). A third gene in the GTF cluster of *Streptococcus salivarius* ATCC 25975. *Aust. Microbiologist* **12**: 344.

STATEMENT OF AUTHORSHIP

The work presented in this thesis is wholly the work of Christine Simpson, except where stated, undertaken at the Institute of Dental Research between March 1990 and March 1994, towards the degree of Doctorate of Philosophy at the University of Sydney.

cd Simpson
30.3.94

ABSTRACT

The glucosyltransferases (GTFs) of oral streptococci polymerize the glucose moieties of sucrose to form glucans. In *Streptococcus salivarius* ATCC 25975, two *gtf* genes, *gtfJ* and *gtfK*, have previously been cloned and sequenced. While it was known that these genes lay next to each other on the *S. salivarius* chromosome and were transcribed in the same direction, it was not known whether they formed an operon. Following the development of a procedure for the isolation of intact RNA from *S. salivarius* ATCC 25975, this possibility was investigated. Both Northern hybridization and RNase protection mapping confirmed that the two genes were transcribed separately and did not form an operon.

It was also known that *S. salivarius* ATCC 25975 produced at least one primer-independent GTF that had not been cloned and sequenced. Following the screening of a gene bank of chromosomal DNA from *S. salivarius* ATCC 25975 constructed in λ L47.1, two new *gtf* genes, *gtfL* and *gtfM*, were isolated. Both of these genes coded for primer-independent GTFs. GtfL synthesized an insoluble glucan, while GtfM synthesized a soluble glucan.

The *gtfL* and *gtfM* genes were sequenced. Putative ribosome-binding sites and -10, -35 promoter elements were identified as were the signal sequence cleavage sites in the deduced amino acid sequences. Both proteins possessed a typical GTF domain structure with a signal sequence, an N-terminal non-conserved domain, a conserved 'catalytic' domain and a C-terminal glucan-binding domain containing a number of 'YG' repeats.

The deduced amino acid sequences of GtfL and GtfM were compared with those of nine other streptococcal Gtfs and a phylogenetic tree constructed. Gtfs with similar products or primer-requirement did not cluster. The GtfL and GtfM were found to be more closely related to the primer-independent GtfD of *S. mutans* than to any other Gtf.

The four GTFs, GtfJ, GtfK, GtfL and GtfM, of *S. salivarius* were found, by ^{13}C -NMR analysis, to synthesize glucans with different structures. The glucans produced by the primer-independent activities, GtfL and GtfM, as well as the primer-dependent GtfK, were found to act as acceptors for primer-dependent glucan synthesis by GtfJ and GtfK.

ABBREVIATIONS

A	absorbance
ANGIS	Australian National Genomic Information Service
APG	<i>p</i> -azidophenyl α -D-glucopyranoside
ATP	adenosine 5'-triphosphate
bp	nucleotide base pairs
BPMs	bacterium-insoluble glucan polymer masses
BTP	1,3-bis[tris-(hydroxymethyl)methylamino]propane
°C	degrees celcius
CAT	chloramphenicol acetyltransferase
cm	centimetre
cpm	counts per minute
C-terminus	carboxyl-terminus
CTP	cytodine 5'-triphosphate
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytodine 5'-triphosphate
dn	of <i>Streptococcus downei</i>
dsDNA	double stranded deoxyribonucleic acid
dUTP	deoxyuridine 5'-triphosphate
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
Fig.	figure
g	gram
g	gravity
GF	sucrose
GTF	general term for glucosyltransferase
Gtf	prefix for cloned glucosyltransferases
GTF-I	glucosyltransferase synthesizing insoluble glucan

GTF-S	glucosyltransferase synthesizing soluble glucan
GTP	guanidine 5'-triphosphate
h	hour
Hz	hertz
IgA	immunoglobulin A
IgG	immunoglobulin G
IPTG	isopropyl β -D-thiogalactosidase
kbp	nucleotide kilobase pairs
$^{\circ}\text{K}$	degrees Kelvin
kJ	kilojoules
K_m	kinetic constant
L	litre
LB	Luria-Bertani
LTA	lipoteichoic acid
M	molar
mg	milligram
MHz	megahertz
min	minute
ml	millilitre
mol	mole
μg	microgram
μl	microlitre
μM	micromolar
mM	millimolar
mRNA	messenger ribonucleic acid
MOPS	3-[N-morpholino]propane-sulphonic acid
M_r	molecular weight
NMR	nuclear magnetic resonance
N-terminus	amino-terminus
PAS	periodic acid Schiff's
PEG	polyethylene glycol
pfu	plaque forming units
pH	degree of acidity

pI	isoelectric point
PIPES	(piperazine-N,N'-bis[2-ethanesulphonic acid])disodium salt
ppm	parts per million
PspA	surface protein of <i>Streptococcus pneumoniae</i>
RBS	ribosome binding site
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal ribonucleic acid
S	substrate concentration
sb	of <i>Streptococcus sobrinus</i>
SDM	semi-defined medium
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	second
Spp	sucrose promoted phenotype
ssDNA	single stranded deoxyribonucleic acid
ssRNA	single stranded ribonucleic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
ToxA	toxin A
ToxB	toxin B
Tris	Trizma base
TTC	2,3,5-triphenyltetrazolium chloride
U	units
UTP	uridine 5'-triphosphate
UV	ultraviolet
V	velocity
V	volts
V _{max}	maximum velocity
v/v	volume per volume
W	watts
w/v	weight per volume
X-gal	5-brom-4-chlor-3-indolyl- β -D-galactopyranoside

Normal chemical nomenclature is used.

CHAPTER 1

GLUCOSYLTRANSFERASES OF ORAL STREPTOCOCCI

The α -D-glucosyltransferases (GTFs) of oral streptococci are large enzymes (approximately 160 kDa) which catalyse the formation of glucans from the glucose moiety of sucrose. These glucans are important in the adhesion and aggregation of oral streptococci and make up most of the extracellular matrix of plaque. GTFs are broadly classified according to two criteria. The first of these is their requirement for dextran as a primer for polymer synthesis and the second, the solubility of their glucan product. Those GTFs that produce soluble products are designated GTF-S while those enzymes that produce an insoluble product are designated GTF-I.

1.1 STRUCTURE OF STREPTOCOCCAL GLUCANS

Glucans are composed of linked α -D-glucosyl residues. Sidebotham (1974) in a review of structural studies concluded that all native dextrans are branched and contain both α -1,6- and α -1,3- linked glucosyl residues, though some contain α -1,2- or α -1,4- branch linkages. These latter linkages seem rare in streptococcal glucans although they have been reported (Sato *et al.*, 1984; Eifuku *et al.*, 1989; Sato and Inoue, 1991).

The types and proportion of linkages determine the physical properties of the glucan (reviewed by Walker 1978). Model building calculations show that glucans with α -1,6 linked glucosyl residues are loosely jointed and thus form flexible chains whereas α -1,3-linked glucosyl residues are stiff and extended and form fibrous aggregates. Determination of the structure of the glucan produced by *Streptococcus sobrinus* OMZ176 in the presence and absence of dextranase demonstrated differences in the physical properties of the native and modified glucans (Ebisu *et al.*, 1974). The insolubility property appeared attributable mainly to the linear α -1,3-linked backbone chain, whereas the α -1,6-linked side-chains were related to the adhesion of the glucan to the surface of the tooth. A more recent study looked at the changes with time in the glucan produced by the

GTF-I of *S. sobrinus* AHT as α -1,3-linked glucosyl residues were added to a main α -1,6-linked chain provided as primer (Takehara *et al.*, 1992). The greater the number of side chains, the sooner the product became insoluble. Thus insolubility depends not only on the proportion of α -1,3-linked glucosyl residues, but also on the degree of branching.

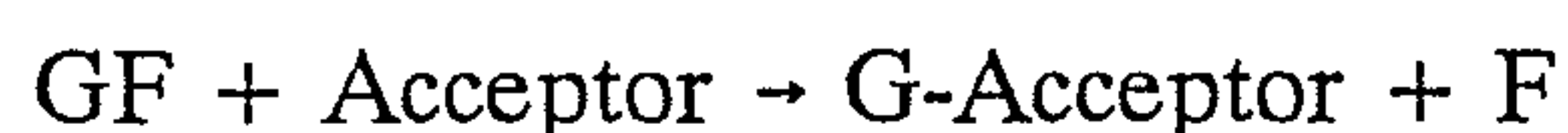
The glucans produced by the GTFs of oral streptococci are diverse. Insoluble glucans produced by oral streptococci contain a backbone of consecutive α -1,3-linked glucosyl residues with α -1,6-linked side-chains (Fukui *et al.*, 1982; Shibata *et al.*, 1983; Ebisu *et al.*, 1974; Davis *et al.*, 1986a; 1986b; Eifuku *et al.*, 1989). These glucans, however, are the result of the combined action of all the GTFs of an organism and do not reflect the properties of glucans produced by individual enzymes.

The enzyme activities which synthesise water-soluble and water-insoluble glucans have been separated into two fractions in a number of studies and the products of these fractions characterised (Hare *et al.*, 1978; Sato *et al.*, 1984; Furuta *et al.*, 1985). Water-soluble glucan has an α -1,6-linked backbone with α -1,3,-branches whereas water-insoluble glucan is composed of linear α -1,3-linked glucosyl residues. Combining the two fractions leads to a highly branched structure containing long chains of both α -1,3- and α -1,6-linked glucosyl residues. Recent studies have indicated that most strains of oral streptococci produce more than two GTFs suggesting that these two enzyme fractions were not pure and that the soluble and insoluble products previously analysed did not necessarily reflect the product of individual enzymes.

1.2 BIOSYNTHESIS OF GLUCANS

The GTFs of oral streptococci are known to catalyse four reactions (Mayer, 1987). These are set out below, where GF is sucrose, G is glucose and F is fructose.

Reaction 1. Glucosyl transfer to an acceptor



Reaction 2. Autopolymerization



Reaction 3. Hydrolysis of sucrose



Reaction 4. Isotope transfer



Sucrose is the only naturally occurring substrate known for these enzymes. The unique oxygen bridge between the anomeric carbon atoms of the glucose and fructose moieties of sucrose enables glucans to be synthesised extracellularly by GTFs without any further input of energy. This is made possible since the free energy of hydrolysis of the α - β bond of sucrose is extremely exothermic (27.6kJ mol⁻¹ at pH 7.0 compared with that of 31.0kJ mol⁻¹ for ATP) (Jacques and Giffard, 1991).

Mutans streptococci are unable to utilise structural isomers of sucrose which generally inhibit glucan synthesis in the presence of sucrose (Minami *et al.*, 1990). Other synthetic donors, however, have been studied, and these include lactulosucrose, glucosyl sorboside and to a limited degree glucose oligosaccharides (Binder *et al.*, 1983). α -1-Fluoroglucose has also been shown to donate glucosyl residues to acceptors, indicating that a glucosidic linkage is not critical for activity (Jung and Mayer, 1981).

An enzyme which does not require sucrose as a substrate has also been described (McCabe and Hamelik, 1983). It catalyses the transfer of radioactivity from donor [¹⁴C]isomaltosaccharides to acceptor dextran in the absence of sucrose forming linkages other than α -1,6-linked glucosyl residues on the acceptor dextran. This enzyme may not be a GTF. Kinetic studies with a GTF of *S. sobrinus* found no evidence of glucosyl transfer between dextrans (Mooser *et al.*, 1985).

1.2.1 Reaction 1: Glucosyl Transfer to an Acceptor

Of the four reactions catalysed by streptococcal GTFs, only the first two lead to the formation of glucans. Reaction 1 depends on the availability of acceptor molecules. Unlike donor substrates, acceptor specificity is extremely broad, although efficiencies vary greatly.

Dextran is thought to serve as an enzyme activator (increasing the velocity of the reactions) and also a primer (its molecular weight increases in the reaction

with sucrose) (Mayer *et al.*, 1981). The early work of McCabe and Smith (1978) suggested that the acceptor acted primarily as a soluble matrix, on which GTF synthesized potentially water-insoluble structures (i.e. more highly branched). They suggested that the soluble, macromolecular acceptor delayed the onset of precipitation (and thus enzyme inactivation) allowing increased synthesis of insoluble glucans.

Further studies on the acceptor specificity of the GTF of *Streptococcus sanguis* indicated that free monosaccharides were poor acceptors, with acceptor activity increasing with polymer size (Mayer *et al.*, 1981). Acceptor activity was found to be significantly increased when glucosyl residues were fixed in a glycosidic bond of an α -configuration (Mayer *et al.*, 1981). Several studies have indicated that a chain length of eight (Germaine *et al.*, 1974; Koga *et al.*, 1983) is the minimum required for stimulation. Lower molecular weight acceptors (maltose and α -methyl glucoside) have limited acceptor activity and in the presence of *S. sanguis* GTF were found to form a series of oligosaccharides which were related both as precursors and products. This indicated that chain growth occurs by a series of single glucosyl transfers (Mayer *et al.*, 1981).

Many studies have concentrated on the role of dextran as an acceptor molecule in order to determine what structural feature stimulated GTF-I activity. For instance, Germaine and Schachtele (1976) tested dextrans of two different molecular weights as primers for *Streptococcus mutans* GTF. Equivalent weight-percent solutions of dextrans of different molecular weight were found to be equally efficient in priming new dextran synthesis. As these primers would have the same concentration of nonreducing ends, but different concentrations of reducing ends this suggested chain growth was from multiple nonreducing termini present in dextran molecules. Modification of reducing ends with sodium borohydrate had no effect on its priming ability, lending support to this argument. Differing K_m values for two different dextrans were used in support of this mechanism (Figures and Edwards, 1981). In contrast, the early work of Robyt and Corrigan (1977) led to the conclusion that non-reducing ends were not required as acceptors since glucans chemically blocked at the nonreducing ends (and with side-chains considered too short to be capable of priming) acted as

acceptors.

Further studies using the GTF-I of *S. mutans* showed that the ability of a linear dextran to stimulate glucan synthesis was similar to a dextran with approximately 5% α -1,3 branch linkages, and thus ten times more non-reducing termini (Walker and Schuerch, 1986). This indicated that it may be sequences of α -1,6-branched glucosyl residues and not end groups which stimulate GTF-I activity. If this were the case, the first step would be the synthesis of a branch linkage on the α -1,6-linked chain. The GTF-I of *S. mutans* is known to convert isomaltosaccharides to branched saccharides by the introduction of a D-glucosyl side chain to the D-glucose residue penultimate to the reducing termini (Walker, 1980). A study of the activation of GTF-I with a wide range of dextrans showed that it was dextrans which had the lowest proportion of non-reducing termini and the highest proportion of unbranched α -1,6-linked sequences which were the most efficient (Walker and Schuerch, 1986). Dextrans already containing some α -1,3-linked sequences in addition to α -1,6-linked sequences were the best acceptors. It has been suggested that the reaction of GTF-I with dextran is slow until a more efficient acceptor is produced at which stage the reaction proceeds more rapidly (McCabe and Smith, 1978; Walker and Schuerch, 1986). Long α -1,3-linked sequences are not effective as acceptors for primer-dependent GTF-Is (Walker and Schuerch, 1986).

After branching it is likely that chain lengthening occurs through the non-reducing ends. Addition of D-glucose to isomaltosaccharides by the GTFs from *S. sobrinus* OMZ 176 was at the non-reducing end (Walker, 1973). When reacted with maltodextrin acceptors, the GTFs of *S. sobrinus* 6715 transferred a glucose residue to carbon-6 of either the non-reducing end or the reducing-end residues of the maltodextrins, forming α -1,6 linkages (Fu and Robyt, 1991). When transfer was to the non-reducing residue the product was also an acceptor resulting in a series of products. In contrast transfer to the reducing end residue resulted in a product that was a poor acceptor.

From the above discussion it is clear that it is not known what defines the branch linkages synthesised by GTFs. However, one study has confirmed that a single enzyme can catalyse bond formations of more than one type. Glucosyl

residues are normally transferred to the carbon-6 of monosaccharide acceptors by the GTF of *S. sanguis*, however sugars blocked at this position also serve as acceptors (Bhattacharjee and Mayer, 1993). The formation of α -1,3 bonds on these acceptors suggests that the acceptor binds to the enzyme in a flipped orientation resulting in the carbon-3 being positioned next to the active site (Bhattacharjee and Mayer, 1993).

1.2.2 Reaction 2: Autopolymerization

The autopolymerization pathway proceeds without the addition of acceptors. Studies with the GTF of *S. sanguis* immobilised on hydroxyapatite demonstrated that the products are formed on the enzymes and remain tightly associated with it (Parnaik *et al.*, 1983). The direction of chain growth is different from that of Reaction 1. The immobilised enzyme treated with a short pulse of radioactive sucrose yields an oligosaccharide labelled at the reducing end. A subsequent chase with cold sucrose decreases the radioactivity at the reducing end as a function of time (Ditson and Mayer, 1984). Thus chain growth occurs at the reducing end. This result supports a mechanism proposed by Ebert and Schenk (1968) for dextran synthesis by *Leuconostoc mesenteroides*, whereby the new glucosyl residues are inserted between the growing chain and the enzyme. It is thought that the protein itself serves as a primer for polymer growth generating its own acceptors for glucosyl transfer.

Little is known about the differences between primer-independent enzymes which are not stimulated by exogenous dextran and primer-dependent enzymes which presumably can carry out either glucosyl transfer to an acceptor or inefficient autopolymerization. There are also differences between enzymes of the same class. For example, a high molecular weight glucan produced by a primer-independent GTF-S of *S. sobrinus* appears suddenly in solution, without evidence of oligosaccharides first being synthesized, suggesting that the product is covalently bound to the enzyme. In contrast, another primer-independent GTF-S of *S. sobrinus*, producing oligosaccharides, releases a number of smaller products as catalysis proceeds (Yamashita *et al.*, 1988a).

1.2.3 Reaction 3: Hydrolysis of sucrose.

Luzio and Mayer (1983) confirmed an earlier observation that GTFs catalyse the hydrolysis of sucrose (Goodman *et al.*, 1955). The hydrolysis of sucrose may be viewed as the transfer of the D-glucosyl group to water. The rate of this reaction is slow but does occur under all sets of conditions (Luzio and Mayer, 1983).

1.2.4 Reaction 4: Isotope Transfer

In the absence of acceptor dextran the enzyme can catalyse isotope exchange reactions between fructose and sucrose (Bourne *et al.*, 1964).

1.3 ISOZYMES OF GLUCOSYLTRANSFERASE

Guggenheim and Newbrun (1969) were the first to isolate different GTF activities from an oral streptococci, isolating three GTFs from mutants *Streptococcus* strain HS. Kuramitsu (1976) later noted that none of the numerous smooth colonial mutants of *S. mutans* GS5 were completely devoid of GTF activity suggesting that the bacterium possessed more than one structural gene for GTF.

Different serotypes of mutans streptococci produce different extracellular glucans, these differences are reflected mainly in the amount of insoluble glucan produced (Trautner *et al.*, 1981; Sato *et al.*, 1984). *S. sobrinus* (serotype *d*) produces the most insoluble glucan. The presence of this serotype in human dental plaque is more often correlated with dental caries than is the presence of any other serotype (Huis in't Veld *et al.*, 1979).

S. sobrinus (serotypes *d* and *g*) produces four GTFs (Takehara *et al.*, 1984), which have been separated and studied extensively by a number of groups. The primer-dependent GTF-I produces a glucan that is essentially composed of 100% α -1,3-linked glucosyl residues (Fukui *et al.*, 1982; Hanada and Takehara, 1987b). The other three enzymes produce soluble glucans of different structure (Taylor *et al.*, 1990). One primer-independent GTF-S produces oligo-isomaltosaccharides which are unable to act as primers for GTF-I (Takehara *et al.*, 1984; Hanada and Takehara, 1987a; Yamashita *et al.*, 1988a; Yamashita *et al.*, 1989b; Taylor *et al.*, 1990). The glucan produced by the other primer-independent GTF-S has chains of α -1,6-linked glucosyl residues with a low percentage (9%) of α -1,3 branch

points (Taylor *et al.*, 1990; Hanada *et al.*, 1993). The glucan synthesized by the primer-dependent GTF-S contains chains of α -1,6-linked glucosyl residues with a higher percentage (32%) of α -1,3-branch points (Shimamura *et al.*, 1982; Taylor *et al.*, 1990; Hanada and Takehara, 1991; Hanada *et al.*, 1993a). The proportions of these four GTFs secreted by *S. sobrinus* in continuous culture are affected by the growth rate of the bacteria, the pH of the medium, the limiting nutrient and the presence of Tween 80 (Walker *et al.*, 1990). The enzymes are immunologically distinct but in some cases cross react with enzymes of *Streptococcus cricetus* (serotype *a*) and *Streptococcus downei* (serotype *h*) (Yamashita *et al.*, 1989b; Tsumori, 1991).

GTFs of other oral streptococci have not been as well characterised, though *S. downei* (serotype *h*) has also been reported to produce four GTFs (Russell *et al.*, 1987; Gilmore *et al.*, 1993).

The view that *S. mutans* (serotype *c*) produces three GTFs has been proven by the recent cloning and sequencing of three *gtf* genes in *S. mutans* GS5 (see section 1.4). The genes code for a primer-dependent GTF-S, a primer-independent GTF-I, and a primer-independent GTF-SI which reportedly produces both soluble and insoluble glucan. The amount of GTF-SI has been shown to vary between strains of *S. mutans* without affecting the adhesion of the bacterium to smooth-surfaces (Fukushima *et al.*, 1993).

Early studies with *Streptococcus salivarius* suggested it produced at least three GTFs, two primer-dependent and one primer-independent activity (Pitty *et al.*, 1989). Currently *Streptococcus sanguis* and *Streptococcus gordonii* are thought to express only a single GTF (Jacques, personal communication).

1.4 CLONING OF GLUCOSYLTRANSFERASE ACTIVITIES

The difficulty in separating GTF activities from culture fluid lies with their ability to form aggregates and to remain active despite proteolytic degradation. These problems together with the low yields obtained when purifying some activities, has led some workers to take a genetic approach to help determine the number and nature of enzymes produced by a given species of *Streptococcus*.

In recent years cloning has confirmed the existence of multiple *gtf* genes

within individual streptococcal species. These *gtf* genes encode for enzymes that produce structurally different glucans. GTF activities express well in *Escherichia coli* and are primarily associated with the cytoplasmic membrane as no activity is secreted into the periplasmic space (Hanada and Kuramitsu, 1988; Aoki *et al.*, 1986). *E. coli* does not utilise and has no transport system for sucrose, but GTF activities are readily detected after cell lysis. Cloning of genes into bacteriophage- λ vectors allows for the initial screening of genes encoding GTF on plates containing sucrose. The formation of polymer in plaques of recombinant bacteriophage expressing GTF activity, or the increased growth of *E. coli* cells around plaques due to the release of fructose by such GTFs, is easily detected on these plates (Russell *et al.*, 1985a). Screening of gene banks constructed in both bacteriophage- λ and in plasmid vectors have also been carried out using specific anti-sera raised against GTF preparations (Pucci *et al.*, 1987; Hanada *et al.*, 1991).

Genes encoding two of the four GTFs of *S. downei* MFe28 have been reported in the literature and the recombinant products analysed (Gilpin *et al.*, 1985; Russell *et al.*, 1985a; Russell *et al.*, 1987; Gilmore *et al.*, 1993). The product of the *gtfI* gene is a primer-dependent GTF-I producing a water-insoluble glucan consisting almost entirely of α -1,3-linked glucosyl residues (Russell *et al.*, 1987). The *gtfS* gene encodes a primer-independent GTF-S producing unbranched α -1,6-linked oligo-isomaltosaccharides, which act as primers for insoluble glucan synthesis by the GtfI of *S. downei* (Russell *et al.*, 1990; Gilmore *et al.*, 1993). Southern blot analysis and biochemical evidence suggest that *S. sobrinus* AHT has very similar GTFs, both at a DNA and a structural level.

A *gtfI* gene similar to that of *S. downei* MFe28 has been cloned from *S. sobrinus* 6715 (serotype *g*) (Abo *et al.*, 1991). A *gtfT* gene coding for a primer-independent GTF-S which produces a high molecular weight glucan has been isolated from *S. sobrinus* OMZ176 (serotype *d*) (Hanada *et al.*, 1991). The glucan is thought to have relatively long α -1,3-linked side chains attached to an α -1,6-linked backbone (Hanada *et al.*, 1993b).

Two activities have been cloned from *S. salivarius* ATCC 25975 (Pitty *et al.*, 1989). The *gtfJ* gene encodes a primer-dependent GTF-I and the *gtfK* gene a primer-dependent GTF-S (Giffard *et al.*, 1991; Giffard *et al.*, 1993). These genes

are contiguous on the *S. salivarius* chromosome.

Three *gtf* genes, *gtfB*, *gtfC* and *gtfD* have been cloned from *S. mutans* (Pucci *et al.*, 1987; Hanada and Kuramitsu, 1988; Hanada and Kuramitsu, 1989). The *gtfB* gene, that codes for a GTF-I activity, has been cloned from both *S. mutans*, GS5 and *S. mutans* MT8148 (Aoki *et al.*, 1986; Fujiwara *et al.*, 1992). The *gtfC* gene has also been isolated from the two *S. mutans* strains, LM7 and GS5, and is located directly downstream from the *gtfB* gene (Pucci *et al.*, 1987; Hanada and Kuramitsu, 1988). The enzyme encoded by the *gtfC* of *S. mutans* LM7 has been reported to be primer-dependent producing a water-soluble glucan (Pucci *et al.*, 1987), whereas that of *S. mutans* GS5 produces both water-soluble and insoluble glucan without primer (Hanada and Kuramitsu, 1988). Interestingly, the GtfB and GtfC of *S. mutans* GS5 have been reported to be primer-independent in *E. coli*, but when expressed in *Streptococcus milleri* are stimulated by primer (Fukushima *et al.*, 1992). In the latter case, the products of GtfB and GtfC formed in the absence of primer, are mainly water-soluble oligosaccharides with the remainder (15-35%) being water-insoluble glucan composed primarily of α -1,3-linked glucosyl residues. In the presence of primer, GtfB and GtfC produce primarily water-insoluble glucan (Fukushima *et al.*, 1992).

An isolate of *S. mutans* was found not to carry both the *gtfB* and *gtfC* genes, but to possess a single gene encoding an enzyme which produced insoluble glucan. This single gene is thought to have originated from a recombination event between the *gtfB* and the *gtfC* genes (Yamashita *et al.*, 1992). This mutant does not demonstrate sucrose-dependent smooth surface colonization.

The *gtfD* of *S. mutans* GS5 codes for a primer-dependent enzyme and produces a soluble glucan (Hanada and Kuramitsu, 1989).

1.5 THE DOMAIN STRUCTURE OF THE GLUCOSYLTRANSFERASES

Of the cloned *gtf* genes, the sequences of nine have been published (Table 1.1). All the genes are closely related to one another and possess the features outlined below (1.5.1-4). Each enzyme has a molecular weight of approximately 160 000 and is highly hydrophilic. From these and other studies (Yamashita *et al.*, 1989a) it has become apparent that the GTFs possess two distinct domains, the 'catalytic'

and glucan-binding domains, that are responsible for sucrase and glucan-binding activities.

Table 1.1: *gtf* genes that have been cloned and sequenced.

Gene	Strain	Enzyme Type	Primer Required	Reference
<i>gtfB</i>	<i>S.mutans</i> GS5	GTF-I	yes/no *	Shiroza <i>et al.</i> , 1987
<i>gtfC</i>	<i>S.mutans</i> GS5	GTF-SI	yes/no *	Ueda <i>et al.</i> , 1988
<i>gtfD</i>	<i>S.mutans</i> GS5	GTF-S	yes	Honda <i>et al.</i> , 1990
<i>gtfI</i> (dn)	<i>S.downei</i> ME28	GTF-I	yes	Ferretti <i>et al.</i> , 1987
<i>gtfS</i>	<i>S.downei</i> ME28	GTF-S	no	Gilmore <i>et al.</i> , 1990
<i>gtfI</i> (sb)	<i>S.sobrinus</i> 6715	GTF-I	yes	Abo <i>et al.</i> , 1991
<i>gtfT</i>	<i>S.sobrinus</i> OMZ176	GTF-S	no	Hanada <i>et al.</i> ,1993b
<i>gtfJ</i>	<i>S.salivarius</i> ATCC25975	GTF-I	yes	Giffard <i>et al.</i> , 1991
<i>gtfK</i>	<i>S.salivarius</i> ATCC25975	GTF-S	yes	Giffard <i>et al.</i> , 1993

* some doubt as to primer-dependence, see 1.4.

1.5.1 The Signal Peptide

The deduced amino acid sequences of all nine *gtf* genes show that the GTFs contain typical streptococcal signal sequences of 35-38 amino acids at their N-termini. The signal sequences of the GTFs are highly conserved. Like all signal sequences these regions consist of a basic N-terminus followed by a hydrophobic region and a polar C-terminal region. N-terminal sequence analysis of the mature form of the *S. downei* GtfI expressed in *E. coli* has shown that a signal sequence cleavage site exists (Ferretti *et al.*, 1987) that conforms with the von Heijne’s ‘-3, -1’ rule (von Heijne, 1983). While GTF signal sequences are recognised in *E. coli*,

the structure of the C-terminus of the protein (see below) appears to prevent secretion through the cytoplasmic membrane (Kato and Kuramitsu, 1991a).

The signal sequence of GtfB has been used to produce a hybrid protein with an α -amylase of *Bacillus* sp. This hybrid protein is secreted in *S. gordonii* (Shiroza and Kuramitsu, 1993) though the level of secretion is poor in *S. mutans* and *S. milleri*.

1.5.2 The 'Variable' Domain

Following the signal sequences in the GTFs is a region of 180-200 amino acids that is highly variable. The significance of this variable region is unknown, though partial removal of this region does not appear to affect enzyme activity (Jacques, unpublished data; Abo *et al.*, 1991). It may be that this region is on the surface of the protein and that selective pressure has maintained its variability to enable the proteins to avoid the immune system of the host (Jacques, personal communication).

1.5.3 The 'Catalytic' Domain

The third region of the GTFs is the so called 'catalytic' domain (Giffard *et al.*, 1991). This region of the enzymes is approximately 900 amino acids in length and is highly conserved. It is in this domain that a conserved aspartic acid residue is found that has been shown to be involved in catalysis (see 1.6 below).

The highly conserved nature of the 'catalytic' domain has been used recently to determine the phylogenetic relatedness of the GTFs (Giffard *et al.*, 1993). This analysis showed that enzymes with similar activities do not generally cluster together. The GtfB and GtfC of *S. mutans* and the GtfIs of *S. downei* and *S. sobrinus* were the only exceptions where evidence of partial recombination between the genes coding for these pairs of GTFs was obtained (Giffard *et al.*, 1993).

1.5.4 The Glucan-Binding Domain

Sequence analysis has revealed that the C-terminal one third of the GTFs contain a number of repeating units, which have been classified as 'A', 'B', 'C', and 'D'

repeats on the basis of sequence similarity (Ferretti *et al.*, 1987; Gilmore *et al.*, 1990; Banas *et al.*, 1990; Abo *et al.*, 1991; Giffard *et al.*, 1991, 1993; Hanada *et al.*, 1993). C-terminal deletion analysis of the GtfB of *S. mutans* has confirmed that this domain is responsible for the binding of GtfB to glucan (Kato and Kuramitsu, 1990). GtfB has a total of six 'A' repeats at the C-terminus and removal of more than three of these results in a marked decrease in enzyme activity. Deletion of all but one repeat results in only sucrase activity, while deletion of the entire region totally inactivates the enzyme. Deletion of the repeat units also prevents the binding of GTF to the cell surface of streptococci (Kato and Kuramitsu, 1991b).

Other studies have also indicated that not all the repeating units are required for glucan-binding and polymer-forming activity. For instance a truncated form of the *S. downei* GtfI containing only the first two of its six 'A' repeats, and neither of its 'B' repeats retains GTF activity as well as the ability to bind glucan (Ferretti *et al.*, 1987). Abo *et al.* (1991) cloned and expressed a region of the *gtfI* of *S. sobrinus* containing four of its six 'A' repeats. The expressed peptide bound α -1,6-glucan but had no sucrase activity (Abo *et al.*, 1991).

In similar biochemical studies, trypsin treatment of GTFs with or without protection with dextran has led to the isolation of 40-60 kDa fractions with affinity for dextran comparable with the native enzymes (Mooser and Wong, 1988; Kobayashi *et al.*, 1989; Wong *et al.*, 1990). These fragments were apparently protected from proteolysis by their tertiary structure. N-terminal sequence analysis of one of the smaller (42 kDa) binding domains from the GTF-I of *S. sobrinus* indicated that it started two-thirds into the 170 kDa enzyme, approximately 60 kDa from the C-terminus consistent with the data obtained from the cloned gene products. Interestingly, glucan protection did not allow the isolation of a glucan-binding domain from a primer-independent GTF-S of *S. sobrinus*, suggesting differences between primer-dependent and primer-independent enzymes (Wong *et al.*, 1990).

Similar repeat units with homology to those in the glucan-binding domain of the GTFs have been found in the C-terminal one-third of the Tox A of

Clostridium difficile (Eichel-Streiber and Sauerborn, 1990). These authors have proposed that the binding of ToxA to cells occurs by way of the C-terminal repeat domain, with the N-terminal domain being responsible for toxic function. Wren (1991) has compared the sequences of the repeat units in the ToxA and ToxB of *C. difficile* with those of four GTFs from *S. mutans* and *S. downei*, as well as those found in the glucan-binding protein of *S. mutans* and in four lytic enzymes from *Streptococcus pneumoniae* and its bacteriophages. His comparison showed that the amino acid sequences of all the 'A' repeats from the glucan-interactive proteins and the repeats of the two clostridial toxins possessed three amino acids that were conserved in almost all cases. These were a tyrosine-phenylalanine dipeptide and a glycine, ten residues upstream of the tyrosine-phenylalanine pair.

Eichel-Streiber *et al.* (1992) also compared the sequences of ToxA and ToxB and the GTFs of the mutans streptococci. A consensus sequence based on 134 repeating units led to the conclusion that these repeats have a modular design which includes a stretch of aromatic amino acids flanked by conserved residues.

A more recent analysis has defined a fundamentally conserved repeating motif, designated the 'YG' motif (Giffard and Jacques, 1994). This twenty-one amino acid repeat contains one or more aromatic residues (of which one is usually tyrosine) followed by a glycine 3-4 residues downstream. This sequence of amino acids were found within, or in close proximity to, the 'A', 'B', 'C', and 'D' classes of repeats as well as the C-termini of the toxins of *C. difficile* and the lysins of *S. pneumoniae*.

The role of the C-terminal domains of GTFs in determining the type of glucan synthesized by a given GTF has not been elucidated. For instance, primer-dependency does not predict the binding characteristics of GTFs to insoluble-dextran affinity columns (Ono *et al.*, 1984). Furthermore, when hybrid proteins were constructed from the *gtfB* and *gtfD* genes of *S. mutans* by swapping the 'catalytic' and glucan-binding domains (Nakano and Kuramitsu, 1992), the activities of these hybrid enzymes were extremely low. Furthermore, the results suggested that the glucan-binding domain alone was not responsible for determining primer-dependency nor the structure of the final glucan product.

1.6 THE CATALYTIC SITE OF THE GLUCOSYLTRANSFERASES

Luzio and Mayer (1983) in their studies on the hydrolysis of sucrose by GTF found that the three activities, polymerization, glucosyl-transfer, and hydrolysis were competitive. They suggested that there was an intermediate common to all reactions and that this might be a D-glucosylated derivative of the enzyme. In subsequent studies, a glucosylated enzyme derivative was immobilised on hydroxyapatite (Parnaik *et al.*, 1983). Two types of saccharide were covalently bound, monomeric glucose and oligosaccharides with an average chain length of seventeen. In the presence of the acceptor, maltose, monomeric D-glucose was transferred to form a trisaccharide. Without acceptor, glucose was released from the complex by hydrolysis, and in the presence of D-fructose, sucrose was formed. These results confirmed that a D-glucosylated form of the enzyme was an intermediate of these reactions.

Subsequent studies by Mooser and coworkers resulted in the identification of an active site amino acid in the GTFs. A covalent glucosyl-enzyme complex was isolated by rapidly denaturing a quenched reaction of *S. sobrinus* GTF-S and radiolabeled sucrose (Mooser and Iwaoka, 1989). Results of alkaline hydrolysis of the denatured glucosyl-enzyme complex and its pH dependency were consistent with a β -glucosyl ester linkage to an active site carboxyl group of either an aspartic or a glutamic acid residue. An active-site peptide was subsequently isolated from the GTF-I and GTF-S of *S. sobrinus* by pepsin digestion of the glucosyl-enzyme complexes (Mooser *et al.*, 1991). The active site sequences were similar and both contained three aspartic acid residues as potential sites of glucose conjugation. Mass spectrometry analysis of partially digested products localised the glucosyl group to Asp-6 within the GTF-I peptide, Asp-Ser-Ile-Arg-Val-Asp-Ala-Val-Asp, which is located approximately 50 kDa from the N-terminus of the GTFs. Site-directed mutagenesis has subsequently confirmed the importance of this aspartic acid residue in the 'catalytic' domain of the GtfB of *S. mutans* GS5 (Asp-451), since its replacement with Glu, Asn and Thr resulted in no detectable sucrase or polymer-forming activity, even though the enzyme still retained dextran-binding activity (Kato *et al.*, 1992). Mutation of nearby aspartic acid residues had no marked effect on activity (Kato *et al.*, 1992).

Regions preceding this active site in the 'catalytic' domain have also been shown to be biologically important, although their function has not been determined. For instance, genetic polymorphisms that do not result in amino acid changes have been noted within a highly conserved region of the amino terminus of the *gtfB* and *gtfC* genes of *S. mutans* (amino acid residues 387-427 of GtfB and 413-453 of GtfC) (Chia *et al.*, 1991; Chia *et al.*, 1993a). Furthermore, a peptide of nineteen amino acids obtained from this region (amino acids 404-427 of GtfB) reversed the inhibitory effect of an antiserum raised against the GtfB and GtfC (Chia *et al.*, 1993a). Interestingly a monoclonal antibody raised against this peptide sequence in GtfC (amino acids 435-453) inhibited the activity of GtfC, but failed to inhibit GtfD, despite the latter enzyme containing the same peptide sequence (Chia *et al.*, 1993b). In another study, antisera to a nearby sequence of fifteen amino acids from GtfB (amino acids 342-356) inhibited the synthesis of water-soluble and water-insoluble glucans by GTFs of *S. mutans* (Dertzbaugh and Macrina, 1990).

Mooser *et al.* (1991) have also suggested that a second amino acid functional group is involved in sucrose catalysis, since a proton donor is needed to facilitate fructose departure. Their view is supported by the earlier photooxidation studies of Koga and Inoue (1981) that suggested the involvement of a histidine residue in GTF catalysis. The GTF of *S. mutans* AHT was completely inactivated under conditions where only histidine residues were modified, suggesting that the imidazole portion of histidine functioned as part of the active site of the GTF.

A recent study by Kobs and Mayer (1991) in which the GTF from *S. sanguis* ATCC 10558 was photolabeled using *p*-azidophenyl α -D-glucopyranoside (APG) has suggested that two acceptor binding sites exist on GTFs. APG acts as an acceptor substrate, but not a donor substrate for GTFs. Photoinactivation with APG was carried out in the presence of D-glucose as an active-site protecting agent, indicating that the specificity of photoinactivation for the APG-binding site of the GTF resides principally with the D-glucosyl moiety. Approximately two molar equivalents of APG per molar equivalent of enzyme were incorporated following photoinactivation. Spectrofluorometry was used to determine the

number of binding sites since the intrinsic fluorescence of the GTF decreases when dextran binds to the enzyme. Two dextran-binding sites were observed for the GTF when fluorescence measurements of the enzyme were determined as a function of dextran concentration.

This finding may support the mechanism suggested by Robyt *et al.* (1974) for polymerization (i.e. reaction 2, Section 1.2) in which he suggested that the enzyme contains two equivalent catalytic sites both of which attack sucrose to give two glucosyl groups covalently linked to the enzyme (Fig. 1.1). The C₆-OH of one of these groups makes a nucleophilic attack on carbon-1 of the other glucose unit forming an α -1,6 bond, releasing a nucleophilic group which can attack another sucrose molecule to give a new glucosyl group. This process continues with the two catalytic groups alternately forming covalent complexes with glucose and dextran. The dextran chain grows by the apparent insertion of the glucose unit between the catalytic group and the reducing end of the polysaccharide.

Another possibility is that branch-forming and chain-elongating activities may have different active and acceptor sites on the GTF (Shimamura *et al.*, 1982). Takehara *et al.* (1991), however, determined that no new branch points were formed after initiation of the reaction of the GTF-I of *S. sobrinus* with primer, with only chain extension occurring thereafter. These results may indicate a single active site on the GTF for both branching and chain lengthening activities.

1.7 GLUCAN BINDING PROTEINS

It has long been known that the addition of dextran to glucose-grown mutans streptococci results in the rapid clumping of these cells (Gibbons and Fitzgerald, 1969). Initially it was thought that the dextran binding sites of the cell-bound GTFs were responsible for this phenomenon. As proteases are common in streptococcal cultures and as trypsin treatment of GTFs can lead to the isolation of peptides with an affinity for dextran (Mooser and Wong, 1988; Kobayashi *et al.*, 1989; Wong *et al.*, 1990) it has been suggested that some putative glucan-binding proteins may in fact be protein fractions resulting from cleavage of GTFs (Mooser and Wong, 1988). However dextran-induced agglutination of *S. mutans* was found to be independent of cell-bound GTF activity (McCabe and Smith, 1975) and

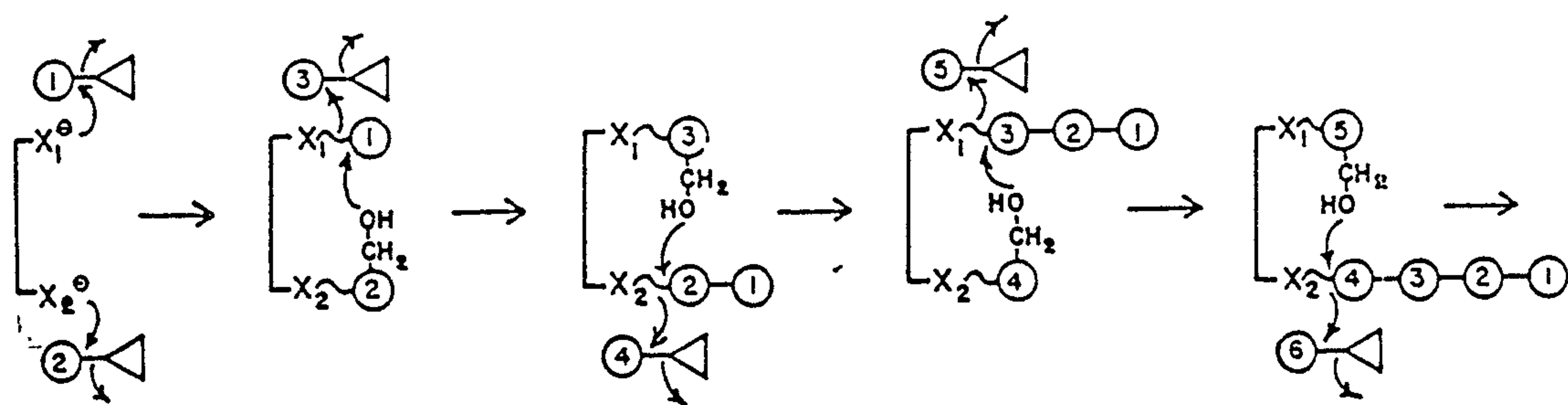


Fig. 1.1. Mechanism for dextran biosynthesis by GTF. X_1 and X_2 represent nucleophilic groups at the active site. $\bigcirc$ represents glucose, $\triangle$ is fructose, $\bigcirc-\triangle$ is sucrose, a $-\bigcirc-\bigcirc-$ is two glucose units linked by an α -1,6-glucosidic bond. The mechanism pictured is concerted, although it could equally well be consecutive.

(Robyt *et al.*, 1974)

competitive inhibition studies suggested that the dextran receptor involved was not identical to that present on the GTF protein (Kelstrup and Funder-Nielsen, 1974).

A study of a variant of *S. mutans* GS5 obtained after continuous culture gave further support for the existence of a glucan-binding site distinct from GTF activity (Janda and Kuramitsu, 1977). The variant was found to be altered in its ability to undergo dextran- and sucrose-induced cell agglutination and to adhere to smooth surfaces in the presence of sucrose, although it produced similar amounts of cell-associated and extracellular GTF activities to the parent strain in chemically defined medium.

McCabe *et al.* (1977) purified a single glucan-binding protein (74 000 molecular weight) from *S. sobrinus* using affinity chromatography. Three proteins which bound glucan were also isolated from *S. mutans* serotype *c* (Russell, 1979). Two of these proteins were GTFs, while the other exhibited no GTF activity and had a molecular weight of 74 000. An antigenically related but non-identical glucan binding protein of molecular weight 87 000 has also been isolated recently from *S. sobrinus* B13 (Wu-Yuan and Gill, 1992). Glucan-binding proteins have also been observed on *Streptococcus cricetus*, *S. sobrinus* and several strains of *S. mutans* (Drake *et al.*, 1988a).

The glucan-binding proteins differ from GTFs in that manganous ion is required for their expression (Drake *et al.*, 1988b). The addition of chelating agents to the culture medium prevents expression of the glucan-binding proteins of *S. cricetus* and *S. mutans*. The presence of chelating agents is also able to (reversibly) inhibit glucan-induced aggregation and sucrose-dependent adhesion of streptococci to glass surfaces (Lu-Lu *et al.*, 1992).

The unique nature of the glucan-binding protein of *S. mutans* Ingbritt was recently confirmed following its cloning and sequencing (Russell *et al.*, 1985b; Banas *et al.*, 1990). The deduced amino acid sequence of this gene showed that the protein possessed a molecular weight of 76 500 and was characterized by a series of repeat sequences which showed homology to the 'A' and 'C' repeat sequences of the GTFs.

A smaller glucan-binding protein has also been isolated from *S. sobrinus* 6715-49. This protein is a homo-dimer of two 7 500 molecular weight sub-units

(Landale and McCabe, 1987). The specificity and affinity for glucans of this glucan-binding protein has been assessed by affinity electrophoresis (Landale and McCabe, 1987). The protein specifically binds to α -1,6-linked glucosyl residues with the highest degree of affinity for isomaltosaccharides possessing a chain length of eight. This was similar to the result obtained by Drake *et al.* (1988a) for the glucan-binding protein of *S. cricetus* and also for the binding of GTFs to dextran (Germaine *et al.*, 1974; Koga *et al.*, 1983). When glucans produced by purified *S. sobrinus* GTFs were tested, the 7 500 glucan binding protein displayed highest affinity for a linear α -1,6-linked glucan produced by primer-independent GTF-S (Landale and McCabe, 1987).

While examining the kinetics of glucan-induced cell aggregation, Drake *et al.* (1988a) determined that only high molecular weight (greater than 5×10^5) glucan with at least 80% α -1,6-linked glucosyl residues promoted strong aggregation. Low molecular weight glucans did not result in aggregation of *S. cricetus* but bound to glucan-binding protein in a competitive manner thus inhibiting higher molecular weight glucan-mediated aggregation.

1.8 REGULATION OF GLUCOSYLTRANSFERASES

1.8.1 Effect of Carbon Source and Growth Rate

Studies have shown variation in the proportions of GTFs produced under various growth conditions in the chemostat (Walker *et al.*, 1990). Production and activity of a primer-dependent GTF-S of strains of *S. sobrinus* is highest at low growth rates, where it accounts for most of the total GTF activity, whereas GTF-I becomes prominent only at high growth rates.

Glucose limitation can also have an effect on the amount of GTF produced although the effect appears to vary for different oral streptococci (Walker *et al.*, 1990). The rate of production of GTF by *S. sobrinus* K1-R was lower under conditions of glucose sufficiency than in cultures where glucose was limiting, regardless of growth rate. For *S. sobrinus* 6715-13-201 there was no evidence of repression of enzyme synthesis. These results differ from previous studies with *S. mutans* strains Ingbritt and 3209 which released more GTF activity in glucose-sufficient medium than in glucose-deficient medium (Wenham *et al.*, 1979,

1981).

GTFs may be extracellular or cell-associated and distribution is affected by sucrose in the growth medium without a change in total GTF activity (Montville *et al.*, 1977). GTF activity in chemically defined media devoid of sucrose is extracellular (Janda and Kuramitsu, 1976; Hamada and Slade, 1979). Addition of sucrose causes this extracellular GTF to become cell-associated (Janda and Kuramitsu, 1978; Hamada and Slade, 1979; Jacques, unpublished observations). This binding of GTF depends upon cell-surface glucan produced in the presence of sucrose (Hamada and Slade, 1979) and can be washed from the surface with sodium dodecyl sulphate (SDS) and low concentrations of Tween 80 (Townsend-Lawman and Bleiweis, 1991).

Recent studies have used chloramphenicol acetyltransferase (CAT) operon fusions to measure the transcriptional activity of the *gtfB* gene promoter (Hudson and Curtis, 1990; Wexler *et al.*, 1993). Transcriptional activity increased with decreasing growth rate in medium containing glucose (Wexler *et al.*, 1993). The addition of sucrose to steady state continuous cultures also resulted in a significant increase in the specific activity of the CAT fusion. However, in some cases, total extracellular GTF activity did not parallel that of the activity of the corresponding CAT fusion (Wexler *et al.*, 1993). It was suggested that this lack of correlation was due to post-transcriptional events controlling enzyme secretion and enzyme activity, as well as a the possible differential expression of dextranase by *S. mutans*.

1.8.2 Effect of Tween 80

Growth of oral streptococci in media containing the non-ionic surfactant Tween 80 can result in enhanced levels of GTF (Umesaki *et al.*, 1977; Wittenberger *et al.*, 1978). Tween 80 has been found to enhance both the production of a GTF-I and two previously undetected primer-independent GTF-Ss by *S. sobrinus* 6715, but to have little effect on the primer-dependent GTF-S of this bacterium (Shimamura, 1983; Walker *et al.*, 1990). Furthermore, the GTF-I of *S. cricetus* could not be detected after growth in media modified by the removal of a high molecular weight fraction which was believed to contain a surfactant(s) that could

stimulate the release of the GTF-I (Walker *et al.*, 1992).

Several workers have suggested that modulation of membrane lipids by Tween 80 may influence GTF secretion and thus its synthesis. Tween 80 stimulates production of extracellular GTF and also increases the degree of unsaturation of the membrane lipid fatty acids in *S. salivarius* ATCC 25975 (Jacques *et al.*, 1985; West *et al.*, 1987). However, growth of *S. salivarius* ATCC 25975 in the presence of n-alkanols and other surfactants also led to the stimulation of GTF production, but failed to increase the level of unsaturated fatty acids in the membrane lipids (Jacques, 1985; Markevics *et al.*, 1987), suggesting that the increase in GTF expression was not simply related to the degree of unsaturation (Pitty and Jacques, 1987). Buchan and Jenkinson (1990) found that for some strains of oral streptococci Tween 80 was inhibitory to GTF production and suggested that the surfactant may have a direct effect on GTF synthesis as well as influencing membrane composition, and that these combined effects might account for the differences observed between different oral streptococci.

1.8.3 Effect of $[Na^+]/[K^+]$

Strains of *S. sanguis* have been reported to require a high $[K^+]/[Na^+]$ ratio for GTF synthesis and secretion (Keevil *et al.*, 1984). Increasing the $[Na^+]$ of the medium decreased GTF production by *S. sanguis* and *S. sobrinus*, but not by *S. salivarius* (Buchan and Jenkinson, 1990). In *S. salivarius* both an increase in $[Na^+]$ or $[K^+]$ enhanced extracellular GTF expression, but in different ways (Markevics and Jacques, 1985; Pitty and Jacques, 1989). The addition of 100mM Na^+ to a defined growth medium devoid of Na^+ initially reduced the rate of expression of extracellular GTF by *S. salivarius* (Pitty and Jacques, 1989). However once *S. salivarius* had adapted to grow in the presence of 100mM Na^+ , the production of extracellular GTF was stimulated. A step up in $[K^+]$ was shown to result only in a transitional increase in extracellular GTF production (Markevics and Jacques, 1985).

1.8.4 Effect of Other Compounds

A number of other compounds found in the oral cavity are known to influence

GTF activity directly. Serum, gingival crevicular fluid, and parotid, submandibular and labial minor gland saliva, stimulate glucan formation by GTFs of *S. sobrinus* 6715 (Schachtele *et al.*, 1978). This effect has been noted in other studies where control and sometimes immune sera stimulate rather than inhibit glucan production (Burckhardt *et al.*, 1976; Evans and Genco, 1973; Fukui *et al.*, 1974; Kuramitsu and Ingersoll, 1976; Linzer and Slade, 1976; Russell *et al.*, 1976). The stimulatory compounds in these fluids are thought to be phospholipids, which are known to stimulate GTF activity (Harlander and Schachtele, 1978). Stimulation is due to interaction between the phospholipid and a site on the enzyme which is distinct from the glucosyl donor or acceptor sites. Thus production of glucans in the oral cavity may be greater than envisioned from *in vitro* studies.

The formation of insoluble glucan by the extracellular GTF of *S. sobrinus* 6715 was found to be stimulated up to 1.6-fold and 2.7-fold in the absence and presence of primer dextran, respectively, by various mono- or di-valent cations (Mukasa *et al.*, 1979). Total GTF activity was stimulated and glucan formation was reported to have changed from soluble to insoluble. Whether there was any effect on the nature of the linkages in the glucan was not reported.

Ammonium ions have also been noted to change the nature of the glucan synthesized by GTFs from soluble to insoluble (Kuramitsu and Wondrack, 1983; Asem *et al.*, 1986; Hanada and Kuramitsu, 1989). In one study, this change was reported to be the result of an increase in the α -1,3-linked glucosyl residues in the resultant glucan (Kuramitsu and Wondrack, 1983). However, in another study the insoluble glucan formed in the presence of NH_4^+ lacked any α -1,3-linked glucosyl residues (Asem *et al.*, 1986).

Hydrogen peroxide is often used with sodium bicarbonate as an antimicrobial agent. Hydrogen peroxide has been found to be a potent inhibitor of the GTF activity of *S. sobrinus* C211 at relatively high concentrations, but to be a stimulator of GTF activity when employed at relatively low concentrations (McAlister *et al.*, 1989). Low concentrations of hydrogen peroxide have been reported to convert the GTF-S of *S. sobrinus* C211 to a GTF-I type enzyme. Nuclear magnetic resonance (NMR) analysis confirmed that the production of α -1,3 linked glucosyl residues increased when the GTF-S was treated with hydrogen

peroxide. The conversion from soluble to insoluble glucan synthesis was considered to be due to aggregation of the enzyme (McAlister *et al.*, 1989b).

1.8.5 Regulatory Element of *S. gordonii* Challis

Little is known about the regulation of GTFs at the molecular level. It is known that the *gtfB* and *gtfC* genes of *S. mutans* GS5 are independently transcribed (Kuramitsu *et al.*, 1991) and recent studies of *S. gordonii* Challis have identified a regulatory gene. In the case of *S. gordonii* Challis, this regulatory element may be unique to this strain which exhibits phase variation between high (Spp⁺, sucrose-promoted phenotype) and low (Spp⁻) levels of extracellular GTF activity (Tardif *et al.*, 1989). Evidence for a regulatory gene comes from a study in which two mutants were isolated from *S. gordonii* Challis which expressed increased levels of GTF (Haisman and Jenkinson, 1991). The mutants expressed 2.6-fold and 4.7-fold more GTF activities than the wild type. A double mutant was constructed and this strain produced 6.4-fold more GTF. Enzymes from wild-type and mutants were biochemically indistinguishable and synthesized structurally identical glucans. The mutations appeared to define separate genetic loci involved in regulating expression of GTF activity in *S. gordonii*.

More recently a gene has been cloned that functionally complements the Spp⁻ phase variant of *S. gordonii* Challis and also positively regulates *gtf* expression (Sulavik *et al.*, 1992). Introduction of a plasmid containing this gene into a Spp⁺ strain resulted in a 10-fold increase in GTF expression. The gene was designated *rgg* and is approximately 1kbp in size. Disruption of the *rgg* reading frame in a Spp⁺ strain resulted in a Spp⁻ phenotype. The gene maps just upstream of the *gtf* of *S. gordonii* Challis.

1.9 CARIES AND COLONIZATION

The adhesion of mutans streptococci to teeth and their subsequent colonization are important in the cariogenic process. The production of glucans from sucrose is a major factor in colonization and smooth surface caries has been shown to be dependent on mutans streptococcal infection and sucrose consumption in animal models (Tanzer 1979). Both rodent and human cariogenic streptococci have been

found to synthesize large quantities of extracellular polysaccharide, particularly from sucrose, whereas non-cariogenic bacteria form only trace amounts (Gibbons *et al.*, 1966). This extracellular polysaccharide is composed entirely of glucose (Gibbons and Banghart, 1967). *S. mutans* has also been observed to adhere to dextran-containing plaques and dextran-treated teeth suggesting that this phenomenon may be of importance in the formation of streptococcal dental plaques.

Production of adherent glucan is not enough to ensure cariogenicity. While *Streptococcus ferus* is able to form plaque like deposits *in vitro* and *in vivo*, infection does not cause dental caries (Freedman *et al.*, 1982). This avirulence appears to result more from the reduced aciduricity of *S. ferus* than from differences in GTF components. The physiology of these *S. ferus* strains demonstrates that in addition to plaque formation, acid tolerance is crucial to the carious process.

1.9.1 Role of Glucosyltransferases in Adhesion

Glucan-synthesis-defective mutants of mutans streptococci are avirulent (Tanzer *et al.*, 1974,) and do not adhere to solid surfaces (Freedman and Tanzer, 1974; Kuramitsu, 1976; Bulkacz and Hill, 1977). Supplementation of a GTF-synthesis-defective mutant with cell-free GTF restores pathogenicity (Hirasawa *et al.*, 1980). Caries also results from a synergistic relation between two avirulent organisms, a GTF mutant and an aggregation-defective mutant, suggesting both properties are required for caries induction (Hirasawa *et al.*, 1980).

Glucan synthesis enables *S. mutans* to adhere to various solid surfaces (reviewed by Hamada and Slade, 1980). Mukasa and Slade (1973; 1974a; 1974b) in a series of studies determined the mechanism of adhesion of *S. mutans* to smooth surfaces. Heat killed cells were unable to adhere to smooth surfaces in the presence of sucrose except with the addition of GTF enzymes. These enzymes were bound to the cell surface. Binding was inhibited by both antidextran globulin and dextranases suggesting that cell-bound dextran formed at least part of the binding site (Mukasa and Slade, 1974a; Hamada and Torii, 1978). This was corroborated by later experiments where cells heat treated after growth in sucrose

were able to bind GTF and adhere to glass surfaces, whereas cells grown in the absence of sucrose could not (Hamada and Slade, 1979).

The adherent properties of mutants of *S. mutans* AHT were found to depend on the proportions of the GTFs produced (Koga and Inoue, 1978). In another study the product of the GTF-I of *S. sobrinus* OMZ176 formed in the presence of dextran T-10, or the product of the GTF-S of *S. sobrinus* in the presence of insoluble glucan were not adherent (Inoue *et al.*, 1982). Glucan produced by the combined actions of the enzymes was, however, adherent, with the amount of insoluble glucan and adherent glucan dependent on the ratios of the enzymes.

The isolation of genes encoding GTFs and the subsequent insertional inactivation has allowed the role of specific GTFs of *S. mutans* to be tested. Both the *gtfB* and *gtfC* gene products which produce insoluble glucan were needed for sucrose-dependent colonization of smooth surfaces. The *gtfD* gene product which produces a soluble product was not required. Inactivation of the *gtfB* and *gtfC* genes resulted in reduced adherence to glass and reduced virulence in a rat model of cariogenicity in which strains were implanted in young gnotobiotic rats fed a 5% sucrose diet (Munro *et al.*, 1991). Additional inactivation of the *gtfD* gene did not reduce virulence any further. A strain of *S. mutans* carrying only a single gene coding for synthesis of insoluble glucan (probably as a result of recombination between the *gtfB* and *gtfC* genes) induced lower levels of smooth surface caries than the wild type stain and could not colonise smooth surfaces *in vitro* in the presence of sucrose (Yamashita *et al.*, 1992). Insertion of the *gtfC* gene into this strain resulted in increased smooth surface caries and sucrose-dependent colonization of smooth surfaces. Inactivation of the *gtfD* gene appeared to have no effect. In another more recent study (Yamashita *et al.*, 1993) the role of each *gtf* gene was determined in a specific-pathogen-free rat model system where rats were fed a diet of 56% sucrose. Inactivation of either or both the *gtfB* and *gtfC* genes resulted in reduced smooth-surface caries, as did inactivation of the *gtfD* gene. Thus while the *gtfD* gene product may not be required for colonization of hard surfaces *in vitro* it may be required for maximal smooth surface caries development when animals are fed a high sucrose diet.

In a similar but far less extensive study, insertional inactivation of the *gtfS* gene, which produces a primer-independent GTF-S in *S. downei* has been shown to prevent the formation of water-insoluble glucan and to reduce the ability of the mutant *Streptococcus* to adhere to glass (Gilmore *et al.*, 1993).

Although it appears that glucan-binding is primarily responsible for cell accumulation in plaque there is some evidence that glucan-binding may also have a role in the initial adhesion to the tooth. GTF is present in saliva and is also incorporated into the initial salivary pellicle that forms on teeth (Scheie and Rolla, 1986). GTF incorporated in the pellicle is active and much of the glucan it produces is retained on the pellicle. *S. mutans* can bind specifically to glucans formed in experimental salivary pellicles as well as host-derived components (Schilling and Bowen, 1992). This suggests that glucan in the salivary pellicle may be a mechanism by which such streptococci can adhere to the tooth surfaces or immature plaque. Glucan on the pellicle also masks host-derived binding sites which may effect the adhesion of those species which do not possess glucan binding proteins.

1.9.2 Glucan as a Diffusion Barrier in Plaque

It has been suggested that the insoluble glucan produced from sucrose may play a role in the initiation of dental caries by acting as a diffusion barrier for the acid produced by fermentation by plaque bacteria (Kleinberg, 1970). Electrodes coated with wild type *S. mutans* showed a decrease in plaque pH to pH 4.0 while growing in a complex medium with a high sucrose concentration, whereas the surrounding medium stayed at pH 6.0 (Hojo *et al.*, 1976). The pH of the plaque produced by a mutant with reduced insoluble glucan-forming ability remained close to that of the medium (Hojo *et al.*, 1976).

More recent work, however, has suggested that this result is not due to reduced acid diffusion through plaque. Several studies have examined the diffusion of sugars through dental plaque and streptococcal cell masses and found no evidence of permselectivity (Dibdin, 1981; McNee *et al.* 1982; Dibdin *et al.*, 1983; Tatevossian, 1985). McNee *et al.* (1982) reported that the diffusion rates of NaF and sucrose appeared to increase as the polysaccharide content in plaque

increased. Other studies have reported that extracellular polysaccharide only has a slight effect on diffusion of carboxylic acids and sugars (Dibdin *et al.*, 1981; Dibdin *et al.*, 1983; Tatevossian, 1985). Tatevossian (1985) reported that with increasing polysaccharide content there was a decrease in sucrose diffusion in dental plaque, but that the opposite occurred with *S. mutans* cell masses.

A recent study with *S. mutans* led to the conclusion that it was increased diffusion of glucose, and thus higher acid production, in cell masses with lower cell density and higher glucan content that led to a lower pH (van Houte *et al.*, 1989). This study, however, did not rule out that prolonged pH depression may still be the result of restricted diffusion of bacterial acids.

1.9.3 Dextran-Mediated Interbacterial Aggregation

Dextran can induce aggregation of other species such as *Actinomyces viscosus* (McBride and Bourgeau 1975; 1975a; Bourgeau and McBride, 1976). Sucrose-grown streptococci will adhere to *A. viscosus* without the addition of dextran as will glucose-grown streptococci in the presence of dextran (Bourgeau and McBride, 1976). Both soluble and insoluble dextrans result in aggregation. This property would allow other species to attach to dental plaque.

1.9.4 GTF Phase Variation of *S. gordonii*

The discovery that *S. gordonii* exhibits phase variation between high (Spp⁺) and low (Spp⁻) GTF activity (Tardiff *et al.*, 1989; see 1.8.5 above) has led to a number of studies concentrating on the ecological implications. Although both phenotypes are able to adhere to hydroxyapatite and saliva-coated hydroxyapatite beads, accumulation of growing cells only occurs with the Spp⁺ phenotype in the presence of sucrose (Vickerman *et al.*, 1991b). This accumulation is associated with both insoluble and soluble glucan formation. The production of glucans by *S. gordonii* Challis does not cause a significant increase in accumulation of the Spp⁻ strain (Vickerman *et al.*, 1991a). This result supports the hypothesis that in the presence of sucrose, Spp⁻ strains are released from the enveloping insoluble glucan mass and that conversely Spp⁺ variants arising from Spp⁻ cell populations utilise available sucrose and accumulate as bacterial glucan masses attached to

surfaces. The release of the less adhesive forms would allow other sites to be colonised. Other changes in phenotype may also occur during phase variation, as Spp⁻ strains show greater adhesion to saliva-coated hydroxyapatite (Vickerman *et al.*, 1992).

Cells of both Spp⁻ and Spp⁺ phenotypes, and GTF-negative cells adhere equally well to preformed bacterium-insoluble glucan polymer masses (BPMs), indicating that attachment to BPMs is independent of cell surface GTF activity (Vickerman and Jones, 1992). Adhesion is inhibited by the presence of sucrose, and by protease treatment of the washed cells. Sucrose releases lipoteichoic acid (LTA) from the BPMs, and adhesion is inhibited by LTA, suggesting LTA may have a role in the binding to BPMs. Addition of dextran had no effect on adhesion of cells to BPMs. This was further evidence that glucans play no role in *S. gordonii* adhesion to BPMs. In support of this notion Hiroi *et al.* (1992) found that *S. gordonii* cells could not attach well to pellicles formed from saliva supplemented with dextrans or GTFs, suggesting that *S. gordonii* cells do not have glucan-binding proteins. Long term incubation with sucrose and with mutans streptococcal GTFs allowed loosely bound *S. gordonii* cells to attach more firmly. This indicates that GTFs have a role to play in accumulation but not adhesion of *S. gordonii* to BPM. These findings are in contrast to those for mutans streptococci, where GTFs and glucan-binding proteins are thought to confer an advantage to this species in adhesion and accumulation in developing plaque (Section 1.9.1).

1.9.5 Sucrose-Independent Adhesion

Glucan synthesis is not the sole determinate of adhesion. A comparison between adhesion of *S. mutans* and *S. sobrinus* to experimental pellicles has shown that *S. mutans* possesses adhesins which bind to salivary components in the pellicle while *S. sobrinus* binds to glucan (Gibbons *et al.*, 1986). Cell-surface receptors to salivary agglutinin allow *S. mutans* to adhere to saliva- or agglutinin-coated early plaque bacteria such as *S. sanguis* and *A. viscosus* (Lamont *et al.*, 1991). The surface-associated protein P1 influences the ability of *S. mutans* to adhere to saliva-coated surfaces and possibly affects primary colonisation of the oral cavity

in the absence of a glucan surface, but has no effect on glucan-mediated adhesion *in vitro* or *in vivo* (Bowen *et al.*, 1991).

1.10 EFFECT OF ANTIPLAQUE AGENTS

A large range of agents have been studied for their effects on glucan synthesis and plaque formation. A possible mode of action for an antiplaque agent would be to reduce the cohesiveness of the plaque matrix glucans.

1.10.1 Glucanases as Antiplaque Agents

Walker (1978) reviewed the effect of dextranases on plaque production and solubilization. Dextranase from *Penicillium funiculosum* has proved to be effective in the removal of plaques produced by some strains of oral streptococci on wires and on hamster's teeth. It was also found to reduce plaque formation in hamsters fed high sucrose diets. However dextranase alone is not always effective in trials with other animals. For human use, dextranases would have to be administered in a form where they were not washed out by saliva possibly through the use of a concanavalin A conjugate which would bind to dextrans (Walker, 1978).

Both dextranases and α -1,3-glucanases have proved effective in reducing glucan synthesis and adhesion. A low concentration of dextranase inhibited GTF and sucrose-dependent adhesion of heat-killed *S. mutans* AHT to glass (Koga and Inoue, 1979) by reducing insoluble glucan formation. It was concluded that dextranase markedly inhibits insoluble glucan production, and reduces adhesiveness of the glucans by lowering their content of α -1,6 linked glucosyl residues and thus their molecular weights.

Insoluble glucan formed on a glass surface from sucrose by the action of cell-free GTFs of *S. sobrinus* OMZ176 was removed by a purified preparation of endo-(1,3)- α -D-glucanase from a strain of *Pseudomonas* sp (Inoue *et al.*, 1988). Synthesis and *de novo* adhesion on glass of the glucan was markedly inhibited by the presence of the α -1,3-glucanase, even at low concentrations. When compared with native glucan, the α -1,3-glucanase-modified glucan contained lower proportions of α -1,3-linkages; showed lower susceptibility to α -1,3-glucanase and higher susceptibility to dextranase; contained a larger amount of low-molecular-

weight fractions; had lower intrinsic viscosities; showed higher *S. sobrinus* cell-agglutinating activities and consisted of looser entwined coalescent single-stranded fibrils and shorter double stranded fibrils (Inoue *et al.*, 1988).

1.10.2 Inhibitors of Glucosyltransferase Activity

A number of antiplaque agents inhibit glucan synthesis because of their structural similarity to sucrose. Structural isomers of sucrose not utilised by mutans streptococci have in most cases been found to inhibit both glucan synthesis and sucrose dependent adhesion to glass surfaces by *S. sobrinus*, but surprisingly not by *S. mutans* (Minami *et al.*, 1990). Carbohydrates with an amino group in the carbon-6 position are potent inhibitors of GTF but only at relatively high concentrations (Thaniyavarn *et al.*, 1981). An exception is acarbose, which has been shown to be a competitive inhibitor of the GTFs from strains of *S. mutans* and *S. sanguis* and to be effective at low concentrations (Newbrun *et al.*, 1983). Maltose and isomaltose are also competitive inhibitors of both GTF-S and GTF-I, whereas nigerose has no effect on GTF-S, but competitively inhibits GTF-I (McAlister *et al.*, 1989a).

Sugar substitutes have also been studied for their effects on glucan production and caries. Sucralose, a non-cariogenic sucrose derivative, inhibits the formation of glucan non-competitively in proportion to the sucralose-to-enzyme ratio but independent of sucrose concentration (Young and Bowen, 1990). A potential sugar substitute, nystose, inhibits GTF activity and adhesion of cells in the presence of sucrose. Lower caries scores are obtained in rats infected with *S. sobrinus* and fed with a diet containing nystose compared with a diet containing sucrose (Ikeda *et al.*, 1990).

Several recent studies have looked at the effect of delmopinol, a highly surface-active compound but not a strong anti-microbial agent. Clinical trials have shown that local applications of delmopinol can reduce plaque formation, delay its maturation and reduce gingivitis in human volunteers (Collaert *et al.*, 1990). Delmopinol can reduce total glucan synthesis as well as reduce the amount of bound glucan (Steinberg *et al.*, 1992). It has little effect on preformed glucans, but reduces the viscosity of glucans synthesized in its presence (Rundegren and

Arnebraunt, 1992). Transmission electron microscopy has shown that plaque is normally dominated by globular or fibrillar matrix components, whereas delmopinol treated plaque possesses empty or unordered matrix areas between more densely packed cells (Rundegran *et al.*, 1992). Delmopinol reduces surface-associated glucan synthesis and lowers the cohesion of the plaque. Thus plaque formed during repeated rinsings with delmopinol is easier to remove by mechanical means than similar non-treated plaque.

Polyphenolic compounds from Oolong tea (Nakahara *et al.*, 1993) and Japanese Green Tea (Otake *et al.*, 1991) inhibit GTFs and have been shown to lower the caries scores for *S. mutans* in pathogen-free rats. Propolis, the resinous hive product collected by bees, which is important in the defence of the hive, similarly inhibits GTF activity and lowers the caries score for *S. sobrinus* in rats (Ikeno *et al.*, 1991).

Several anionic and cationic surfactants, including SDS and cetylpyridinium chloride, have also been shown to enhance or inhibit glucan synthesis and sucrose-dependent cell adhesion to smooth surfaces depending on the concentrations of the surfactants used (Kawabata *et al.*, 1993).

1.11 A ROLE FOR GLUCAN IN ENDOCARDITIS

The initial event of bacterial endocarditis is attachment of bacteria to heart valves (Angrist and Oka, 1963). Dextran production by streptococci has been suggested to be of importance in this initial adhesion. Adhesion to normal aortic valves is low irrespective of whether or not streptococci are grown in media containing sucrose. However, glucan-producing strains of streptococci grown in sucrose are five times more likely to adhere to damaged aortic valves than streptococci grown in sucrose-free medium (Ramirez-Ronda, 1978). Dextranase treatment results in a decrease in the adhesion of these bacteria to levels seen with streptococci grown in medium lacking sucrose. Increased adhesion is restored in the presence of exogenous dextran.

Recently a mutant of *S. mutans* which had been genetically engineered to be devoid of any GTF activity was compared with the wild-type strain (Munro and Macrina, 1993). Rats infected with the wild-type strain developed endocarditis

more frequently than rats infected with the mutant. Similar differences were found between the wild-type strain grown on sucrose and the wild-type strain grown on glucose suggesting that although sucrose is not found in the bloodstream, the polysaccharide carried by the infecting organism is sufficient to cause virulence. Although the wild type strain coated in glucan associated with human granulocytes in similar numbers to cells without glucan it had a greater survival rate in phagocytosis assays. This suggests that the presence of polymer prevents internalization of *S. mutans* cells by granulocytes, an hypothesis supported by microscopic examination (Munro and Macrina, 1993).

1.12 GLUCOSYLTRANSFERASES AS POTENTIAL IMMUNOGENS

If colonization of mutans streptococci could be prevented or reduced this would result in reduced dental caries. One means of achieving this could be the development of a vaccine against *S. mutans* (as reviewed Krasse *et al.*, 1987). GTFs are theoretically good candidates for a vaccine as they are important in promoting adhesion and do not appear to elicit an auto-immune response.

Antibodies to GTFs inhibit GTF activity *in vitro*. Although studies in monkeys have not been encouraging (Russell and Colman, 1981) other animal models have indicated a reduction in caries after immunisation with GTFs (McGhee and Michalek, 1981). Recent human trials have indicated that oral immunisation with GTF in conjunction with an aluminium based adjuvant can elicit a short term salivary IgA response which can interfere with re-accumulation of *S. mutans* after dental prophylaxis (Smith and Taubman, 1987; 1990).

Different GTFs possess different antigenic determinants. However some antibodies have been shown to inhibit a number of GTF activities (Tsumori, 1991). Recently synthetic peptides derived from the repeating sequence of the glucan-binding region of the GtfI of *S. downei* has been shown to elicit high serum IgG response after subcutaneous injection into Sprague-Dawley rats (Smith *et al.*, 1993). These antisera also reacted with intact GTFs from *S. sobrinus* and *S. mutans* and with a glucan binding protein from *S. sobrinus*.

Antisera to peptides derived from the N-terminal 'catalytic' domain of the GTFs have also been shown to inhibit GTF activity (see 1.6 above). For instance,

monoclonal antibody to a peptide of nineteen amino acids based on the sequence of the amino acids 435-453 of the GtfC of *S. mutans* inhibited GtfC enzyme activity. However this monoclonal failed to inhibit GtfD activity, despite the fact that the enzyme contained the same peptide sequence (Chia *et al.*, 1993b). Antisera to a sequence of fifteen amino acids from GtfB (amino acids 342-356) was able to inhibit the synthesis of water-soluble and water-insoluble glucans (Dertzbaugh and Macrina, 1990). However, antisera to active-site peptides, containing the conserved aspartic acid residue, did not result in significant inhibition of the GtfI or the GtfS of *S. downei*, indicating that either the antibodies could not access the active-site region or alternatively did not recognise the native enzyme conformation (Cope and Mooser, 1993).

Antibodies to GTFs or sub-sequence peptides could also modify colonization of the oral cavity by mutans streptococci in ways other than by simple enzyme inhibition. For instance, an antipeptide antibody that is reactive with epitopes on GTFs or glucan binding proteins could agglutinate the mutans streptococci resulting in the clearance of these cariogenic microorganisms from the oral cavity. Similarly these antibodies could inhibit adhesion and thus colonization and accumulation in plaque. Modification of the structure of the glucan synthesized in the presence of antipeptide antibody might also affect the plaque-forming ability of the mutans streptococci or the diffusibility of compounds within plaque.

CHAPTER 2

ROUTINE METHODS

This chapter describes methods used routinely during this work. Experimental methods are described in the relevant chapters as are any 'kits' that were supplied by various manufacturers.

2.1. REAGENTS, CHEMICALS AND ENZYMES

All chemicals were of analytical reagent grade except for those indicated below:

D-glucose (laboratory grade): AJAX Chemicals.

Hydroxylapatite (Bio-Gel HTP): Bio-Rad Laboratories.

Ampicillin, 5-brom-4-chlor-3-indolyl- β -D-galactopyranoside (X-gal), cesium chloride, deoxyribonuclease (DNase) I, lysozyme, proteinase K, ribonuclease (RNase) A, RNA molecular weight marker II: Boehringer Mannheim.

p-Aminobenzoic acid, biotin, 1,3-bis[tris-(hydroxymethyl)methylamino]propane (BTP), L-cysteine, folic acid, guanine, DL-methionine, calcium pantothenate, pyridoxamine-2HCl, riboflavin, thiamine-HCl: Calbiochem-Behring Corp.

Tryptone, yeast extract: Difco Laboratories.

[α -³²P] 2'-deoxyadenosine 5'-triphosphate (dATP), [α -³²P] cytidine 5'-triphosphate (CTP), [α -³²P] 2'-deoxycytidine 5'-triphosphate (dCTP), [α -³⁵S]dATP, [U-¹⁴C-glucosyl]-labelled sucrose and Formula-989: Dupont NEN.

Electrophoresis gel stain: Gradipure Ltd.

Redistilled phenol, xylene cyanol FF: International Biotechnologies, Inc.

Bacteriological agar, casein hydrolysate: Oxoid.

Agarose NA, bind-silane, dextran sulphate, dextran T10, isopropyl β -D-thiogalactosidase (IPTG), repel-silane, DRigestII DNA molecular weight markers, High Molecular Weight - SDS calibration kit, Isoelectric Focusing calibration kit pH 3-10: Pharmacia.

Klenow DNA polymerase, RQ1 RNase free DNase, T3 RNA polymerase, T7 RNA polymerase: Promega Corporation.

Nicotinamide: Roche.

Acrylamide, adenine, adenosine 5'-triphosphate (ATP) (sodium salt), ammonium persulphate, bromophenol blue, diethylpyrocarbonate (DEPC), dithiothreitol (DTT), ethidium bromide, formaldehyde, L-histidine, β -mercaptoethanol, N-laurylsarcosine, N,N'-methylene-bis-acrylamide, 3-[N-morpholino]propane-sulphonic acid (MOPS), periodic acid, (piperazine-N,N'-bis[2-ethanesulphonic acid])disodium salt (PIPES), polyethylene glycol (PEG), RNase T₁, Schiff's reagent, sodium azide, sodium fluoride, sodium dodecyl sulphate (SDS), sucrose, N,N,N',N'-tetramethylethylenediamine (TEMED), thiomersal, 2,3,5-triphenyltetrazolium chloride (TTC), Triton X-100, Trizma base (Tris), Tween 80, uracil: Sigma.

Chaetomium gracile α -1,6-dextranase: Miles Biotechnology Group.

Restriction enzymes, *Hind*III digested bacteriophage- λ molecular weight markers, Calf intestinal alkaline phosphatase and T4 ligase were obtained from either Boehringer Mannheim, International Biotechnologies Inc., New England Biolabs or Pharmacia.

Oligonucleotides were synthesized on a Gene Assembler Plus (Pharmacia LKB), using Pharmacia chemicals, and mutanolysin was kindly supplied by Dr K. Yokogawa, Dainippon Pharmaceutical Co., Osaka, Japan.

2.2 BACTERIAL STRAINS, PHAGES AND PHAGEMIDS

Strains and vectors used in these studies are listed in Table 2.1.

2.3 MEDIA AND GROWTH CONDITIONS

S. salivarius ATCC 25975 was grown at 37°C in semi-defined medium (SDM; Jacques, 1984) with 25mM glucose as the fermentable carbon source (see Appendix 1 for method of preparation). When appropriate, Tween 80 was added to 0.05% (v/v) to maximise GTF production (Pitty and Jacques, 1987).

Escherichia coli strains were grown at 37°C in Luria-Bertani (LB) medium (5g/l yeast extract, 10g/l tryptone, 5g/l NaCl), supplemented with ampicillin (100 µg/ml), or '2xYT' medium (10g/l yeast extract, 10g/l tryptone, 5g/l NaCl; Sambrook *et al.*, 1989) as appropriate. When cloning into pIBI vectors, IPTG (1mM) and X-gal (100µg/ml) were added where appropriate for blue/white colony differentiation.

Liquid lysates of various bacteriophage-λ were prepared by inoculating an appropriate number of aliquots of 50ml or 125ml of LB medium containing 10mM MgSO₄ in 250ml or 1L conical flasks respectively with 2% (v/v) of an overnight culture of *E. coli* LE392 also grown in LB medium supplemented with 10mM MgSO₄. For the 50ml cultures, 10⁸, and for the 125ml cultures, 10⁹, plaque forming units (pfu) of bacteriophage-λ were added. The cultures were shaken at 37-39°C until they lysed. Lysis was completed by the addition of 2% (v/v) chloroform to the culture followed by shaking for a further 5-10min. If the lysate was to be used for DNA preparations, NaCl was added to a concentration of 1M to promote dissociation of phage from bacterial debris.

Table 2.1. Bacterial Strains, Phages and Phagemids

Bacteria, Phage or Phagemid	Description	Source or reference
Bacterium:		
<i>Streptococcus salivarius</i> ATCC 25975		ATCC (Hamilton, 1967).
<i>Escherichia coli</i> LE392	F ⁻ e14-(mcrA ⁻) <i>hsdR514</i> (r _K m _K ⁻) <i>supE44 supF58 lacY1</i> or Δ (<i>lacIZY</i>)6 <i>galK2 galT22 metB1 trpR55</i>	Murray <i>et al.</i> , 1977
<i>Escherichia coli</i> JM105	F ['] <i>traD36 lacI^HΔ(lacZ) M15 proA⁺B⁺/thi rpsL</i> (Str ⁺) <i>endA sbcB15 sbcC?</i> <i>hsdR4</i> (r _K m ⁺) Δ (<i>lacproAB</i>)	Yannisch-Perron <i>et al.</i> , 1985
<i>Escherichia coli</i> NM522	F ['] <i>lacI^HΔ(lacZ) M15 proA⁺B⁺/supE thi Δ(lac-proAB) Δ(hsdMS-mcrB)5</i> (r _K m _K ⁻ McrBC ⁻)	Gough and Murray, 1983
Bacteriophage:		
λL47.1		Loenen and Brammar, 1980
λA-8	λL47.1 with GtfJ encoding 8.5kbp <i>Sau3A</i> partial fragment of <i>S. salivarius</i> ATCC 25975	Pitty <i>et al.</i> , 1989
λA-33	λL47.1 with GtfK encoding 9.6kbp <i>Sau3A</i> partial fragment of <i>S. salivarius</i> ATCC 25975	Pitty <i>et al.</i> , 1989
λC-13	λL47.1 with 8.3kbp GtfM encoding <i>Sau3A</i> partial fragment of <i>S. salivarius</i> ATCC 25975	This study
λD-10	λL47.1 with 11kbp GtfL encoding <i>Sau3A</i> partial fragment of <i>S. salivarius</i> ATCC 25975	This study
λD-40	λL47.1 with <i>Sau3A</i> partial fragment of <i>S. salivarius</i> ATCC 25975 isolated from sucrose-containing medium	This study
λB-5.2	λL47.1 with 8.3kbp <i>Sau3A</i> partial fragment of <i>S. salivarius</i> ATCC 25975 containing 0.8kbp <i>EcoRI/SacI</i> of λA-8	This study
M13K07		IBI Corporation

Table 2.1. Bacterial Strains, Phages and Phagemids

Bacteria, Phage or Phagemid	Description	Source or reference
Phagemid:		
pIBI30	Ap ^r , f1 origin replication, β-galactosidase, T3 and T7 polymerase promoters	IBI Corporation
pIBI31	Ap ^r , f1 origin replication, β-galactosidase, T3 and T7 polymerase promoters	IBI Corporation
pGSG101	pIBI30 with GtfJ encoding 6.8kbp <i>SacI/Bam</i> HI fragment of λA-8	Giffard <i>et al.</i> , 1991
pGSG103	pIBI30 with 0.9kbp <i>Eco</i> RI fragment of λA-8	Giffard <i>et al.</i> , 1991
pGSG104	pIBI30 with 0.8kbp <i>Eco</i> RI/ <i>Sac</i> I fragment of λA-8	This study
pGSG105	pIBI30 with 2.3kbp <i>Eco</i> RI fragment of λA-8	This study
pGSG201	pIBI30 with GtfK encoding 7.3kbp <i>Bgl</i> II/ <i>Bam</i> HI fragment of λA-33	Giffard <i>et al.</i> , 1991
pGSG202	pIBI30 with 1.85kbp <i>Eco</i> RI fragment of λA-33	This study
pGSG203	pIBI30 with 3.0kbp <i>Eco</i> RI/ <i>Sac</i> I fragment of λA-33	This study
pGSG301	pIBI30 with 0.8kbp <i>Sac</i> I/ <i>Eco</i> RI fragment of λB-5-2	This study
pGSG302	pIBI30 with 1.6kbp <i>Sac</i> I/ <i>Hind</i> III fragment of λB-5-2	This study
pGSG303	pIBI30 with 3.8kbp <i>Sac</i> I/ <i>Eco</i> RI fragment of λB-5-2	This study
pGSG304	pIBI30 with 4.6kbp <i>Eco</i> RI fragment of λB-5-2	This study
pGSG305	pIBI31 with 1.6kbp <i>Sac</i> I/ <i>Hind</i> III fragment of λB-5-2	This study
pGSG306	pIBI30 with 6.5kbp <i>Sac</i> I/ <i>Bam</i> HI fragment of λB-5-2	This study
pGSG401	pIBI30 with GtfL encoding 6.2kbp <i>Bam</i> HI/ <i>Xba</i> I fragment of λD-10	This study
pGSG402	pIBI31 with 6.2kbp <i>Bam</i> HI/ <i>Xba</i> I fragment of λD-10	This study
pGSG403	pIBI30 with 4.8kbp <i>Eco</i> RI fragment of λD-10	This study
pGSG404	pIBI30 with 4.1kbp <i>Eco</i> RI fragment of λD-10	This study
pGSG501	pIBI31 with GtfM encoding 5.3kbp <i>Xba</i> I fragment of λC-13	This study
pGSG502	pIBI31 with GtfM encoding 4.8kbp <i>Eco</i> RI/ <i>Xba</i> I fragment of λC-13	This study
pGSG503	pIBI31 with 3.7kbp <i>Kpn</i> I/ <i>Xba</i> I fragment of λC-13	This study

2.4 DNA MANIPULATIONS

2.4.1 Small Scale Purification and Isolation of Bacteriophage- λ DNA

Small scale isolation of bacteriophage- λ DNA followed the procedure described in 'Promega Protocols'. Essentially, a 20ml lysate was centrifuged (12 000g, 10min, 4°C) and the supernatant incubated with 1 μ g/ml RNase A and 1 μ g/ml DNase I for 30min at 37°C. Following the addition of 20ml of 50mM Tris-HCl (pH 8.0) containing 20% (w/v) PEG, 2M NaCl, 8mM MgSO₄ and 0.01% (w/v) gelatin, the solution was left on ice for 1h. The precipitant was recovered by centrifugation (16 000g, 20min, 4°C) and resuspended in 1ml of phage buffer (20mM Tris-HCl pH 7.4, 100mM NaCl, 10mM MgSO₄). This was followed by sequential extractions with an equal volume of Tris-HCl-saturated phenol (pH 8.0), an equal volume of phenol:chloroform:isoamylalcohol (25:24:1; v/v/v), and an equal volume of chloroform:isoamylalcohol (24:1; v/v). DNA was precipitated at -70°C for 20min by the addition of an equal volume of isopropanol. The precipitate was collected by centrifugation (16 000g, 5min, 22°C) and resuspended in 50 μ l TE buffer (10mM Tris-HCl pH 8.0, 1 mM EDTA).

2.4.2 Large Scale Isolation and Purification of Bacteriophage- λ DNA

Bacteriophage- λ DNA was isolated from 1L liquid lysates according to the method of Silhavy *et al.* (1984) as outlined below. After centrifugation (6000g, 10min, 4°C), 10% (w/v) PEG was added to the supernatant at room temperature (22°C) and the resulting suspension stirred until the PEG was completely dissolved. Bacteriophage- λ was co-precipitated with PEG by placing the solution on ice for 60min. The pellet was collected by centrifugation (6000g, 10min, 4°C) and resuspended in 5-10ml of TM buffer (50mM Tris-HCl pH 7.5, 10mM MgSO₄). The PEG-phage suspension was extracted with an equal volume of chloroform, the aqueous layer recovered, and the PEG-chloroform and interface back extracted with an equal volume of TM buffer. The combined aqueous layers were centrifuged (125 000g, 60min, 4°C) through a step gradient formed by overlaying 6ml of 5% (v/v) glycerol on 6ml of 40% (v/v) glycerol. The pellet was resuspended in 1ml of TM buffer, RNase A and DNase I added to a final

concentration of 10 μ g/ml and 1 μ g/ml respectively, and the solution incubated at 37°C for 30min. STEP buffer (50mM Tris pH7.5, 0.4M EDTA, 0.5% [w/v] SDS, 1mg/ml proteinase K) was added to a 0.2 x final volume and incubated at 50°C for 15min. The DNA solution was then sequentially extracted with an equal volumes of Tris-HCl-saturated phenol, phenol:chloroform:isoamylalcohol (25:24:1; v/v/v) and chloroform:isoamylalcohol (24:1; v/v), and the DNA precipitated with 2 volumes of ethanol, spooled onto a pasteur pipette and resuspended in 1ml of TE buffer (pH 8.0). The resulting DNA suspension was drop-dialysed on a Millipore Type VS filter, with a pore size of 0.025 μ m, against TE buffer (pH 8.0) before use.

2.4.3 Rapid Isolation and Purification of Phagemid DNA

Phagemid DNA was isolated from *E. coli* by the method of Silhavy *et al.* (1984). A single colony was inoculated into 5ml of LB medium containing the appropriate antibiotics (e.g. 100 μ g/ml ampicillin) and shaken at 37°C for 16h. The overnight culture was centrifuged (12 000g, 4°C, 5min), the cell pellet resuspended in 200 μ l of solution I (50mM glucose, 25mM Tris-HCl pH 8.0, 10mM EDTA pH8.0, 5mg/ml lysozyme) and the resulting suspension incubated at room temperature (22°C) for 5min, before addition of 400 μ l of solution II (0.2M NaOH, 1% [w/v] SDS) and further incubation at room temperature (22°C) for 5min. Following the addition of 0.3ml of solution III, the mixture was frozen at -70°C for 10min. Solution III was prepared by adding 60ml of 5M potassium acetate to 11.5ml glacial acetic acid and 28.5ml H₂O. After thawing, cell debris was pelleted (16000g, 22°C, 15min) in a microfuge. Nucleic acid was precipitated by addition of isopropanol to the supernatant to a 0.6 x final volume, and the precipitate recovered by centrifugation (16 000g, 22°C, 5min). The pellet was rinsed with 70% (v/v) ethanol and dried under vacuum, before resuspending in 50-100 μ l of TE buffer (pH 8.0).

2.4.4 Large Scale Isolation and Purification of Phagemid DNA

Phagemids were purified from *E. coli* by the alkaline lysis method essentially as described by Maniatis *et al.* (1982). *E. coli* was grown overnight at 37°C with

vigorous shaking in 500ml LB medium with the appropriate antibiotic (e.g. 100µg/ml ampicillin). Cells were harvested by centrifugation (4000g, 4°C, 15min), resuspended in 10ml solution I (2.4.3 above) and incubated at room temperature (22°C) for 5min. Solution II (20ml; 2.4.3 above) was added to the suspension, mixed and left to stand on ice for 10min. Subsequently, 15ml of solution III (2.4.3 above) was added and the resulting suspension incubated on ice for a further 10min. Bacterial debris was removed by centrifugation (48,000g, 4°C, 30min) and the DNA precipitated by the addition of two volumes of ethanol. The DNA was recovered by centrifugation (12 000g, 22°C, 30min) and dissolved in 16ml of TE buffer (pH 8.0). Cesium chloride was added to 1g/ml before the addition of 1.6ml of ethidium bromide solution (10mg/ml in H₂O). The cesium chloride solution was transferred to an ultracentrifuge tube and the remainder of the tube filled with light paraffin oil. After centrifugation (200 000g, 20°C, 40h), the lower band consisting of closed circular phagemid DNA was removed with a hypodermic needle and syringe. The ethidium bromide was extracted with equal volumes of H₂O-saturated 1-butanol and the resulting aqueous phase dialyzed against several changes of TE buffer (pH 8.0). The DNA was then precipitated by the addition of 0.1 volume of 3M sodium acetate (pH 4.8) and 2.5 volumes of ethanol. The ethanol precipitant was stored at 4°C until required, when it was recovered by centrifugation (16 000g, 22°C, 5min) and resuspended in TE buffer (pH 8.0).

2.4.5 Preparation of Single Stranded Phagemid DNA

For the preparation of single stranded DNA (ssDNA), an overnight culture of *E. coli* (200µl) containing an appropriate phagemid (Table 1) was inoculated into 10ml of '2xYT' medium containing 60µl of ampicillin (100mg/ml) and grown with shaking at 37°C to mid-log phase. A portion of this culture (2ml) was then transferred to a fresh flask, bacteriophage M13KO7 added to a concentration of 2×10^8 to 4×10^8 pfu/ml and the resulting suspension shaken at 37°C for 1h. Subsequently, a portion of this second culture (400µl) was transferred to 10ml of '2xYT' medium supplemented with 60µl of ampicillin (100mg/ml) and 14µl kanamycin (50mg/ml), and shaken at 37°C for 16h. Following centrifugation (12 000g, 4°C, 10min) 1ml of 25% (w/v) PEG in 3M NaCl was added to the

supernatant and the resulting solution left at room temperature (22°C) for 15min before further centrifugation (12 000g, 4°C, 30min). The pellet was resuspended in 400µl of 20mM Tris-HCl (pH 7.5) containing 20mM NaCl and 1mM EDTA, and extracted twice with Tris-HCl-saturated phenol and twice with chloroform:isoamyl alcohol (24:1; v/v). The DNA, precipitated by the addition of a 0.1 volume of 3M sodium acetate and 2.5 volumes of ethanol at -70°C for 30min, was collected by centrifugation (16 000, 22°C, 5min), dried under vacuum and resuspended in 50µl of TE buffer (pH 7.5). The presence of phagemid ssDNA was verified by electrophoresis of an 8µl sample on a 0.7% (w/v) agarose gel (see 2.4.9 below), followed by staining with ethidium bromide (0.5µg/ml).

2.4.6 Spectrophotometric Determination of Nucleic Acid Concentration

The nucleic acid concentration was determined spectrophotometrically (Sambrook *et al.*, 1989). A sample of DNA (5µl) was diluted in TE buffer (pH 8.0) to a volume of 1ml and the absorbance of the resulting solution obtained at 260nm and 280nm in a Beckman DU-64 spectrophotometer or in a Pharmacia GeneQuant DNA/RNA Calculator. In the latter case readings were also taken at 320nm. An A_{260} of 1.000 corresponded to approximately 50µg/ml for dsDNA and 40µg/ml ssDNA or RNA. The A_{260}/A_{280} ratio gave an estimate of the purity of the sample as pure preparations of DNA and RNA have a ratio of 1.8 and 2.0 respectively.

2.4.7 Precipitation of DNA by Ethanol

DNA was precipitated by the addition of a 0.1 volume of 3M sodium acetate and 2.5 volumes of ethanol at -70°C for 5 to 30min. The DNA was recovered by centrifugation (16 000g, 22°C, 5min) in a microfuge and dried under vacuum.

2.4.8 Digestion of DNA with Restriction Endonucleases

Restriction endonuclease digestion of DNA was routinely carried out in 50µl volumes, using the restriction buffer and conditions supplied by the manufacturer. Incubation times varied between 1h and 20h depending on the restriction endonuclease. Heat inactivation at 80-85°C for 10min was routinely used to inactivate the restriction endonuclease prior to the digestion fragments being

analysed by gel electrophoresis (2.4.9 below).

2.4.9 Gel Electrophoresis of DNA on Agarose Gels

DNA was separated on 0.7% (w/v) agarose gels prepared in TBE buffer (89mM Tris, 89mM boric acid, 2mM EDTA pH 8.0) containing ethidium bromide (0.5 µg/ml), using either the 50ml 'mini-gel' or 150ml 'maxi-gel' system manufactured by Hoefer. Electrophoresis was carried out in TBE buffer containing ethidium bromide (0.5 µg/ml) at either 100-150V for 1-4h or 30V for 16h. Samples were diluted to 8µl in TE buffer (pH 8.0) to which 2µl of loading dye (0.25% [w/v] bromophenol blue, 0.25% [w/v] xylene cyanol FF, 30% [v/v] glycerol) was added. *Hind*III digested bacteriophage-λ was generally used as molecular weight standards, except when smaller sized fragments were to be detected, when DRigestII markers (Pharmacia) were used. Gels were routinely photographed on a UV transilluminator (Ultra-Violet Products, Inc.) with a polaroid camera fitted with a Kodak Wratten 23A orange filter using type 667 Polaroid film.

2.4.10 Dephosphorylation

Self-ligation of vector DNA was prevented by dephosphorylation using calf intestinal alkaline phosphatase. The restriction digest was precipitated with ethanol (2.4.7. above) and resuspended in 50µl of 1 x alkaline phosphatase buffer (supplied by the manufacturer). Alkaline phosphatase (1µl, 1-2U) was added and the solution incubated at 37°C for 60min. The reaction was stopped by the addition of 5mM EDTA followed by heating at 75°C for 10min. After extraction with Tris-HCl-saturated phenol and chloroform:isoamyl alcohol (24:1; v/v), the DNA was recovered by ethanol precipitation (2.4.7 above).

2.4.11 Recovery of DNA from Agarose Gels

DNA was recovered from agarose gels using the method of Weichenban (1991). Essentially, a square piece of Whatman 3MM blotting paper was wrapped around the blunt end of a micropipette tip to form a 'beaker' and this then fitted into a microfuge tube which had been punctured at the bottom. The agarose slice was placed into the 'blotting paper beaker' and the tube mounted into a second

microfuge tube. Following centrifugation (16 000g, 22°C, 1min) in a microfuge, the buffer from the agarose slice containing the DNA was retrieved from the bottom tube and extracted with an equal volume of water-saturated butanol to remove ethidium bromide.

2.4.12 Ligation

DNA fragments that were to be ligated were precipitated with ethanol (2.4.7 above) and resuspended in a total volume of 50µl of 1 x ligation buffer (supplied by the manufacturer). T4 ligase (1µl, 5-7U) was then added and the solution left at 4°C for 16h. The ligated DNA was subsequently precipitated with ethanol (2.4.7 above) and resuspended in 10µl of H₂O or TE buffer (pH 8.0).

2.4.13 Preparation of Cells for Electroporation

Cells that were to be electroporated were prepared by a method derived from the instruction manual for the 'Gene-Pulser' unit (BioRad). One litre of LB medium was inoculated with a 0.01 x final volume of an overnight culture of the relevant strain of *E. coli* grown at 37°C in the same medium. This culture was then grown at 37°C with vigorous shaking to an A₆₀₀ of 0.500 - 1.000, at which time the culture was chilled to 4°C on ice (15 to 30min) before centrifugation (4000g, 4°C, 15min). The pelleted cells were washed first in 1L of 1mM BTP buffer (pH 7.0), followed by 20ml of 10% (v/v) glycerol (3000g, 4°C, 15min) prior to resuspending to a final volume of 2 - 3ml in 10% glycerol at 4°C. This suspension was divided into 40µl aliquots and stored at -70°C until required.

2.4.14 Transformation by Electroporation

An aliquot of cells that had been prepared for electroporation (2.4.13) was thawed on ice, 2µl of DNA from the ligation (2.4.12) added and the mixture of cells and DNA transferred to a cold (0°C) electroporation cuvette. The 'Gene Pulser' (BioRad) was set at 25µF and the 'Pulse Controller' (BioRad) to 200ohm. A pulse of 2.5kV was applied to those cuvettes with a 0.2cm electrode gap and 1.6kV to cuvettes with a 0.1cm electrode gap. LB medium (1ml) was immediately added and the cell suspension transferred to a microfuge tube and incubated at

37°C for 1h before plating on selective medium.

2.4.15 Transfer of DNA to Nylon Membranes for Southern Hybridizations

DNA cleaved with an appropriate restriction endonuclease(s) was separated by electrophoresis on a 0.7% (w/v) agarose gel at 30V for 16h. The gel was photographed and molecular weight standards marked onto an aligned Hybond N+ nylon membrane (Amersham). The gel was then washed sequentially in 0.5M NaOH containing 1.5M NaCl for 1h, 0.5M Tris-HCl (pH 7.5) containing 1.5M NaCl for 1h and 10 x SSC buffer (1.5M NaCl, 0.15M trisodium citrate, pH 7.0) for 5min before transferring onto the Hybond N+ nylon membrane. The transfer was accomplished by capillary blotting (2.4.16).

Following the transfer, the DNA was fixed onto the membrane either by exposure to UV radiation for 5min on a transilluminator (Ultra-Violet Products, Inc.), or by placing the membrane for 20min on a filter paper pad soaked in 0.4M NaOH followed by washing in 2 x SSC buffer. Membranes were covered in plastic wrap and stored at 4°C until required (0 - 16h).

2.4.16 Capillary Blotting of Agarose Gels

Essentially a pad of six pieces of 3MM paper (Whatman) in a plastic box were soaked in 10 x SSC buffer (2.4.15) and the prepared gel centred on this pad and surrounded by plastic wrap. The aligned Hybond N+ membrane was placed on the gel ensuring there were no air bubbles. Three pieces of 3MM paper the size of the gel were soaked in 10 x SSC buffer and placed over the membrane. A stack of tissues was placed on top of the 3MM paper and these in turn held in place by a glass plate weighed down with a mass of 0.5 - 1.0kg. Transfer of DNA from the gel to the membrane was allowed to take place overnight (16h).

2.4.17 Hybridizations

Prehybridizations and hybridizations were carried out using either a Hybrid-Ease Hybridization Chamber (Hoefer) in a water-bath set at the desired temperature, or a Hybaid Midi-Oven. Labelling of probes, hybridization buffers and temperature conditions were as described in Chapters 3, 4 and 5. Stringency

washes are also described in the relevant experimental chapters and were carried out in either a plastic box in a water bath, or in the Hybaid Midi-Oven.

2.4.18 DNA Sequencing

Sequence determinations were carried out using ssDNA (Section 2.4.5) or cesium chloride purified dsDNA (Section 2.4.4) by the chain-termination method of Sanger *et al.* (1977). All sequencing reactions used the Pharmacia T7 sequencing kit according to the manufacturer's instructions using T7, T3 or custom-synthesized oligonucleotide primers and [³⁵S]-dATP (Dupont NEN).

Sequencing reactions were run using a Hoefer Poker Face I Nucleic Acid Sequencer. Sequencing gels (8% [w/v] acrylamide, 28% [w/v] urea) were pre-run at 110W for at least one hour. Three loadings were applied to the gel, at 0h, 2h, and 3.5h, with a total running time of approximately 4.5h. After loading, the gel was run at 10W until the loading dye moved into the gel (approximately 5min), then the power was increased to 110W. Gels were fixed with a solution containing 10% [v/v] methanol and 10% [v/v] glacial acetic acid before drying and subjecting to autoradiography.

2.5 DETECTION OF GLUCOSYLTRANSFERASE ACTIVITY

2.5.1 Detection of Sucrase Activity

Sucrase activity was detected using a qualitative microtitre reducing-sugar test similar to that described by Aduse-Opoku *et al.* (1989). A loopful of cells were suspended in 100μl of 50mM sodium phosphate buffer (pH 6.5) containing 0.1% (w/v) SDS and incubated at room temperature (22°C) for 15min. Alternatively 100μl of bacteriophage lysate was used (2.3). To this was added 100μl of 50mM sodium phosphate buffer (pH 6.5) containing 1% (w/v) Triton X-100 and 1% (w/v) sucrose. Dextran T-10 (0.02%, w/v) was also added when assaying primer-dependent or unknown activities. Incubation was continued at 37°C for 1h. Following the addition of 100μl of a solution of triphenyltetrazolium chloride (0.1%, w/v) in 2M NaOH, the mixture was incubated at 37°C for a further 15min. The development of a red colour indicated the presence of released reducing

sugars.

2.5.2 Detection of Glucan Synthesis

GTF activity was quantified by the incorporation of [U-¹⁴C-glucosyl]-labelled sucrose into glucans (Jacques, 1985). GTF activity encoded by phagemids was released from *E. coli* cells by permeabilizing 1ml of a stationary phase culture grown overnight (16h) in LB medium containing ampicillin (100µg/ml). This was achieved by vortexing the cells in the presence of a 50µl solution of SDS (0.1%, w/v) and 100µl of chloroform for 20s. GTF activity was measured at 37°C in 100mM potassium phosphate buffer (pH 6.0) containing 10mM NaF, 0.05% (v/v) Tween 80, 0.1mM histidine and 10mM [U-¹⁴C-glucosyl]-labelled sucrose in a total volume of 2ml. To assay primer-dependent GTF activities dextran T-10 was added to 0.02% (w/v). After incubation for 60min, total glucan was collected by precipitation with 3 volumes of ethanol and the assay tube washed with 5ml of 75% (v/v) ethanol. The ethanol precipitate was trapped by filtration through 2.5cm Whatman GF/B glass fibre filters under vacuum. The filter was washed three times with 5ml volumes of 75% (v/v) ethanol. Alternatively water-insoluble glucan was trapped on the filters and washed with 100mM potassium phosphate buffer (pH 6.0). The dried filter disks were counted in a Beckman Liquid Scintillation Counter (model LS-9000) using Formula-989 as the scintillant.

2.5.3 SDS-PAGE Gel Electrophoresis for the Detection of Glucosyltransferase Activity

SDS-PAGE gels were run on a Hoefer SE 600 system according to the method of Laemmli (1970) as modified by Russell (1976). Samples (100µl) were solubilized at room temperature (22°C) in 35µl of SDS sample buffer (0.2M Tris, 40% [v/v] glycerol, 4% [w/v] SDS, 20% [v/v] β-mercaptoethanol, 0.001% [w/v] bromophenol blue) for 90min before applying to the gel. Molecular weight markers (10µl of High Molecular Weight-SDS calibration kit) were similarly treated prior to application to the gel. Gels were run at 15mA until the tracking dye passed through the stacking gel when the current was increased to 30mA. Alternatively, gels were run at 60V overnight for 16h. The lanes containing

molecular weight markers were removed, fixed in 10% (w/v) trichloroacetic acid for 30min and stained with 'Gradipure' stain overnight, before destaining in 6% (v/v) glacial acetic acid. The GTF activity present in the sample lanes was detected as described below (2.5.4).

2.5.4 Detection of Glucosyltransferase Activity on SDS-PAGE Gels

GTF activity on SDS-PAGE gels was determined by a periodic acid Schiff (PAS) reaction (Pitty *et al.*, 1989). After 3 x 30min washes in 50mM Tris-HCl buffer (pH 7.5), the SDS-PAGE gels were incubated at 37°C for 1-16h in 50mM sodium phosphate buffer (pH 6.5) containing 1% (w/v) sucrose, 1% (v/v) Triton X-100, 0.01% (w/v) thiomersal and 0.02% dextran T-10. PAS staining was then carried out following the fixation of the gel in 75% (v/v) ethanol for 30min. Essentially, the gel was shaken for 60min in a 5% (v/v) glacial acetic acid solution containing 0.7% (w/v) periodic acid followed by a further 3 x 20min washes in a 5% (v/v) glacial acetic acid solution containing 0.2% (w/v) sodium metabisulphite. Following these washes, the gel was placed in Schiff's reagent for 60min, followed by 3 x 30min washes in a 40% (v/v) methanol solution containing 0.5% (w/v) potassium metabisulphite and 5% (v/v) acetic acid. The stained gel was then destained by repeated washes in a 40% methanol solution containing 5% (v/v) glacial acetic acid.

CHAPTER 3

ISOLATION OF RNA AND TRANSCRIPTIONAL ANALYSIS OF THE *gtfJK* LOCUS

3.1 INTRODUCTION

The *gtfJ* and *gtfK* genes of *S. salivarius* ATCC 25975 that encode primer-dependent GTFs have recently been cloned and sequenced (Giffard *et al.*, 1991, 1993). Restriction maps of the inserts of the two recombinant bacteriophages, λ A-8 and λ A-33, carrying these genes, indicated an apparently overlapping region. This overlap was confirmed by Southern blot analysis (Giffard *et al.*, 1991), and sequencing revealed that the two genes were transcribed in the same direction and separated by only 210bp on the *S. salivarius* chromosome (Fig. 3.1) (Giffard *et al.*, 1991).

The addition of the surfactant Tween 80 to batch cultures of *S. salivarius* ATCC 25975 has been shown to result in a several fold increase in two protein bands associated with GTF activity. No similar increase in any other extracellular proteins was noted (Pitty and Jacques, 1987). This observation was consistent with the suggestion that the GTFs of *S. salivarius* were subject to co-ordinate regulation in response to the presence of Tween 80 in the medium. It was therefore of interest to determine whether the *gtfJ* and *gtfK* genes formed an operon or whether they were transcribed separately.

The isolation of RNA from Gram-positive bacteria of a quality suitable for transcriptional analysis can present a technical challenge as the structure of the cell wall may make it difficult to lyse cells rapidly enough to avoid RNA degradation. This is a particular problem with oral streptococci due to their resistance to lysis by lysozyme. A procedure was therefore devised for the isolation of high quality RNA from *S. salivarius* ATCC 25975, in order that Northern hybridization and RNase protection experiments could be carried out. The successful isolation of mRNA from *S. salivarius* ATCC 25975 enabled transcriptional analysis of the *gtfJ* and *gtfK* genes to be undertaken.

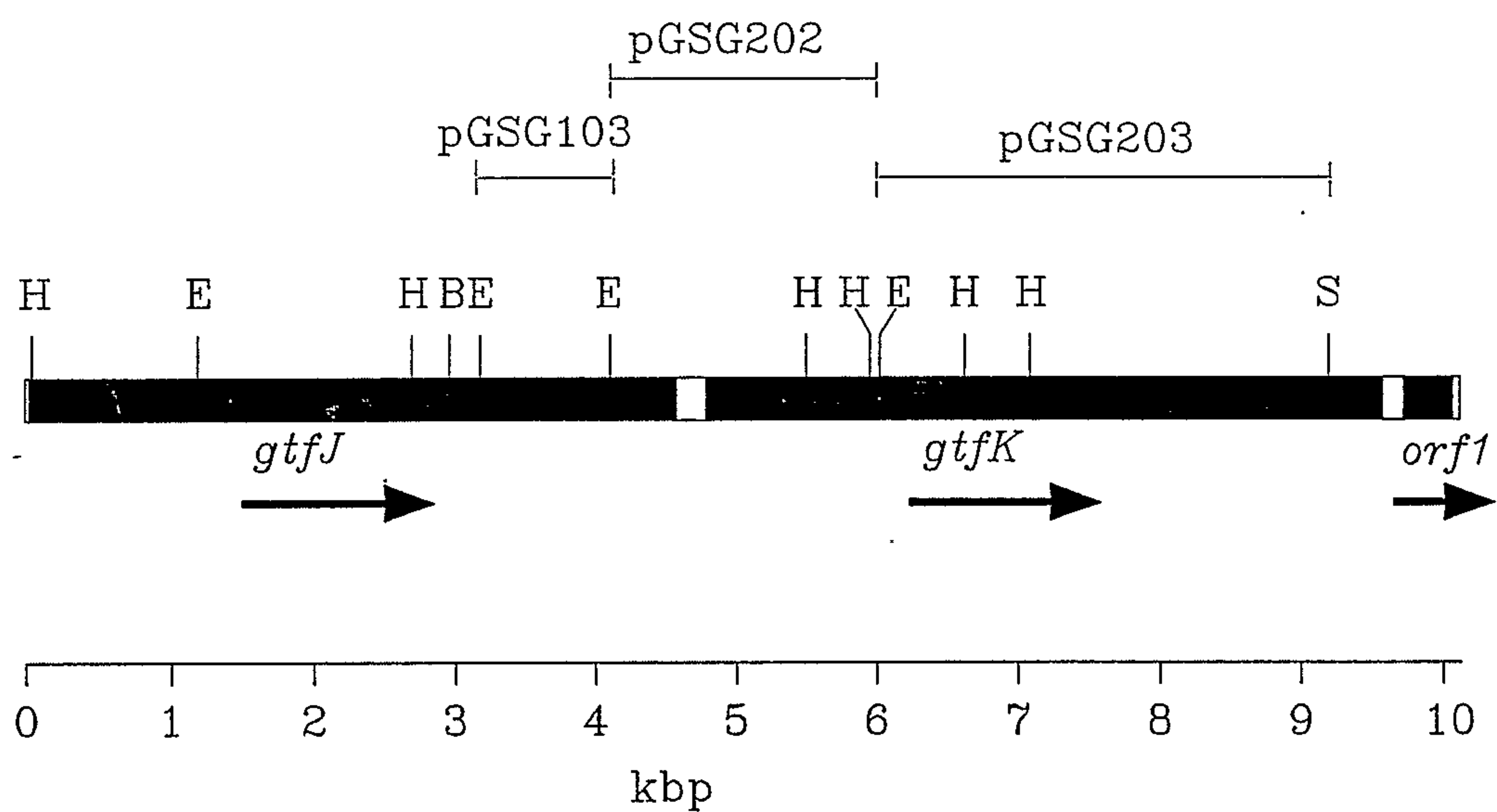


Fig. 3.1. The *gtfJK* locus of *S. salivarius* ATCC 25975 showing the coding regions, directions of transcription, and the fragments used as probes in the experiments described in the text. Restriction site abbreviations are: E, *EcoR*1; H, *Hind*III; B, *Bgl*II; S, *Sac*I.

3.2 MATERIALS AND METHODS

3.2.1 RNA Isolation from *S. salivarius* ATCC 25975

The procedure developed for the isolation of RNA from *S. salivarius* was:

1. Grow *S. salivarius* cells overnight at 37°C in 5ml of SDM (Appendix 1) containing 25mM glucose as carbon source.
2. Dilute the overnight culture in 100ml SDM containing 0.05% (v/v) Tween 80 and grow at 37°C to an A_{600} of 0.100 (Tween 80 was included to maximise GTF expression).
3. Add glycine to 20mM to inhibit cell wall crosslinking and grow at 37°C to A_{600} of 0.600.
4. Centrifuge (4000g, 4°C, 10min) over 100ml of frozen sodium phosphate buffer (pH 6.5) containing 10mM $MgSO_4$ and 1mM $CaCl_2$.
5. Resuspend the bacteria in 2ml of the same buffer, add 200 μ l of 700mM DTT and 200 μ l of aqueous mutanolysin (20mg/ml) and incubate at 60°C for 5min.
6. Add 300 μ l of 10% (w/v) SDS containing 1mM EDTA and boil for 5min.
7. Shear the DNA by repeated passage of the lysate through a medium gauge needle.
8. Add 6ml of 4M LiCl and keep at 0°C overnight.
9. Centrifuge (12 000g, 4°C, 10min) and resuspend pellet in 300 μ l of 100mM sodium acetate containing 0.5% (w/v) SDS.
10. Extract three times with equal volumes of Tris-HCl-saturated phenol (pH 8.0) and once with an equal volume of chloroform:isoamyl alcohol (24:1; v/v).
11. Precipitate RNA by adding 2.5 volumes of ethanol and keeping at -70°C for 10min.
12. Centrifuge (16 000g, 22°C, 5min) in a microfuge and resuspend the RNA in 50 μ l 0.5% (w/v) SDS.

All solutions and plasticware were treated with 0.1% (v/v) DEPC at 37°C overnight to destroy any latent RNase activity and then autoclaved at 121°C for 15min. For the same reason, all glassware was baked at 250°C overnight.

3.2.2 RNA Gel Electrophoresis

RNA was separated by electrophoresis using formaldehyde agarose gels essentially as described by Sambrook *et al.* (1989). The RNA sample (20µg) was mixed with 15µl of running buffer (0.02M MOPS, 8mM sodium acetate, 1mM EDTA) containing 50% (v/v) formamide and 17.9% (v/v) formaldehyde solution (37% in H₂O; v/v), and 2µl of gel loading dye (50% [v/v] glycerol, 1mM EDTA pH 8.0, 0.25% [w/v] bromophenol blue, 0.25% [w/v] xylene cyanol FF) and denatured by heating at 100°C for 5min and then placing on ice. RNA molecular weight markers (Boehringer Mannheim) were prepared in the same manner. Samples were electrophoresed in running buffer, on 1% agarose gels containing running buffer and 2.2M formaldehyde. Gels were run at 35V for 16h, or at 120V for 3-4h. Gels were stained briefly in ethidium bromide (0.5µg/ml) and destained in water.

3.2.3 Northern Hybridizations

After separation by electrophoresis, the RNA was transferred from the gel to a Hybond N+ nylon membrane by overnight capillary blotting in 10 x SSC buffer (2.4.16). The membrane was fixed according to the manufacturer's instructions by placing on a wad of filter paper soaked in 0.05M NaOH for 5min prior to a brief wash in 2 x SSC buffer. Prehybridization (4h) and hybridization (16h) were at 68°C in 0.5M sodium phosphate buffer (pH 7.2) containing 1mM EDTA and 7% (w/v) SDS (as described by Church and Gilbert, 1984) to which was added dextran sulphate to 10% (w/v). The stringency washes were 2 x SSC buffer containing 0.1% (w/v) SDS at room temperature for 10min, followed by 1 x SSC buffer containing 0.1% (w/v) SDS at 68°C for 15min and a final wash of 0.1 x SSC buffer containing 0.1% (w/v) SDS at 68°C for 20min. Hybridizations and washes were carried out using a 'Hybaid Midi Oven' (Hybaid). Probes made from pGSG103 and pGSG203 (Fig. 3.1) were labelled with [α -³²P]dCTP (Dupont NEN) using the Promega 'Prime-a-Gene' random-priming kit according to the manufacturer's instructions.

3.2.4 RNase Protection Assays

³²P-RNA antisense probe was synthesised using a Promega 'Riboprobe System II' kit which makes use of the T7 RNA polymerase. The phagemid template, pGSG202 (Fig. 3.1) was first linearised at a *Bam*HI site situated in the vector at the end of the *S. salivarius* 1.85 kb *Eco*RI insert opposite to the T7 promoter to ensure that only the insert was labelled. Linearised DNA (2µg) was transcribed using 40U of T7 RNA polymerase in 20µl of Promega Transcription Optimized Buffer (40mM Tris-HCl pH7.5, 6mM MgCl₂, 2mM spermidine, 10mM NaCl) containing 20U of RNasin Ribonuclease Inhibitor (Promega), 10mM DTT, 500µM ATP, 500µM GTP, 500µM UTP, 12µM CTP and 50µCi [α -³²P]CTP (Dupont NEN) at 37°C for 60min. The DNA template was removed by digestion with 2U of RQ1 RNase-Free DNase (Promega) at 37°C for 15min, followed by phenol:chloroform:isoamyl alcohol (25:24:1; v/v/v) and chloroform:isoamyl alcohol (24:1; v/v) extractions. Following the addition of 0.5 volumes of 7.5M ammonium acetate and 2.5 volumes of ethanol, the RNA probe was precipitated at -70°C for 30min to remove unincorporated nucleotides. After centrifugation (16 000g, 22°C, 5min), the RNA was resuspended in 100µl of 1M ammonium acetate, re-precipitated as above and dried under vacuum.

Transcription mapping was carried out as described in the Promega Technical Manual for 'Transcription *in vitro* Systems'. RNA was precipitated with ethanol as described above and dried under vacuum before being resuspended with the RNA probe in 30µl of 80%-formamide hybridization buffer (80% [v/v] formamide, 40mM PIPES [pH 6.7], 0.4M NaCl, 1mM EDTA). The sample was heated to 85°C for 5min to denature the RNAs, and then hybridized at 45°C for 6h. After cooling to room temperature, ssRNA was digested by the addition of 300µl of a 'Digestion Solution' which consisted of 10mM Tris-HCl (pH 7.5) containing 5mM EDTA, 300mM NaCl, 2µg/ml RNase T1 and 40µg/ml RNase A. This left the resistant ³²P-RNA/RNA duplex intact. Digestion was terminated by incubation of the solution at 37°C for 30min with 20µl of 10% (w/v) SDS containing 10µl of proteinase K (10mg/ml). The sample was then extracted with phenol:chloroform:isoamyl alcohol (25:24:1; v/v/v), and the RNA precipitated at -20°C for 30min following the addition of 1ml of ethanol. The precipitated RNA

was then centrifuged (16 000g, 22°C, 5min) and the pellet dried under vacuum before resuspending in 10µl of H₂O.

The RNA was subsequently electrophoresed as described above (3.2.2). Following the removal of the lane containing the RNA molecular weight markers and its staining with ethidium bromide, the remaining experimental RNA was transferred from the gel onto Hybond N+ membrane by capillary blotting in 10 x SSC buffer (2.4.16) overnight (16h) and an autoradiograph obtained.

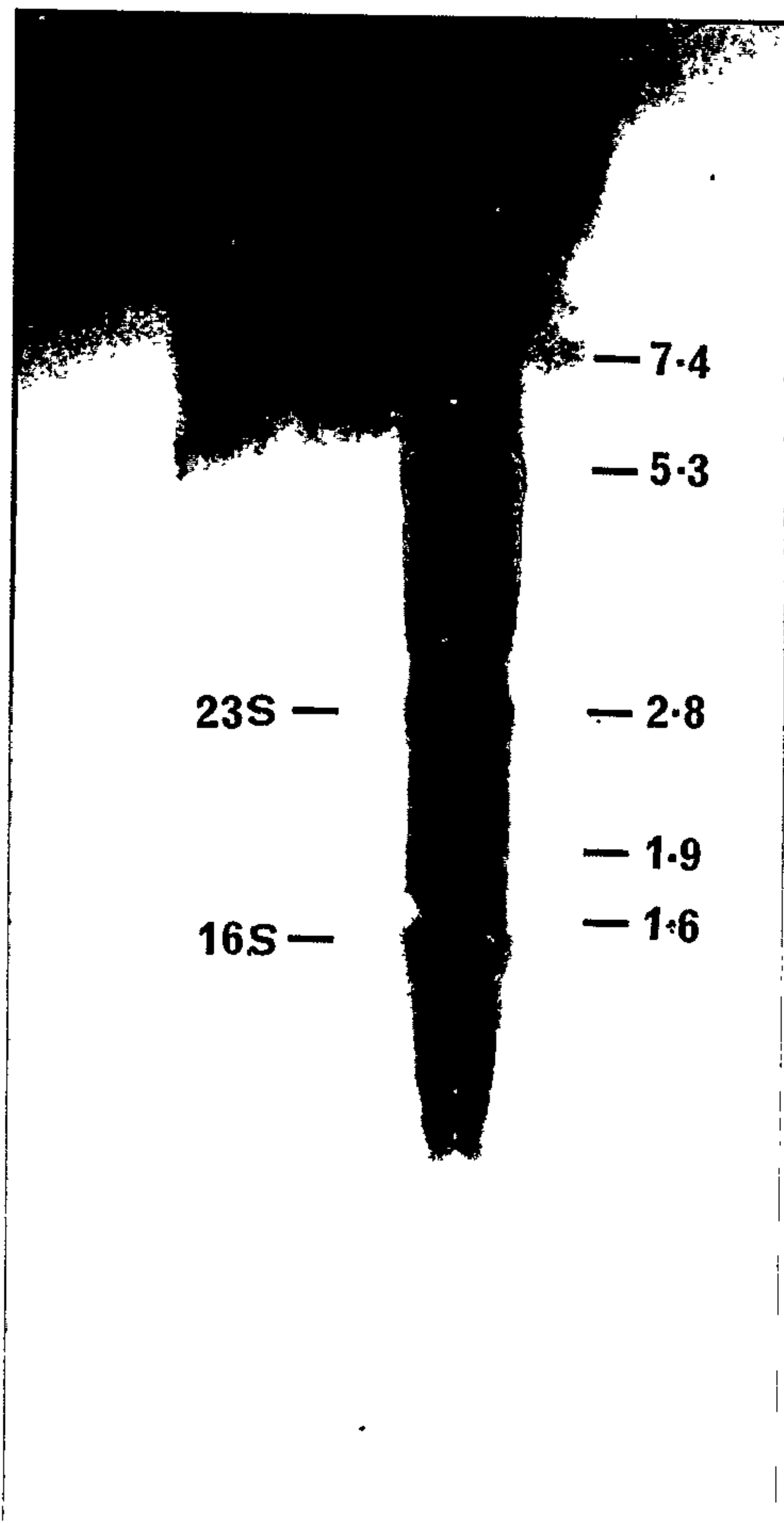
3.3 RESULTS

3.3.1 RNA Isolation

Chromosomal DNA has been isolated from oral streptococci after cell lysis at 37°C with the muramidase, mutanolysin, in either phosphate buffer (Giffard *et al.*, 1991) or Tris-HCl buffer (Ganeshkumar *et al.*, 1988). These procedures require lysis times of 1h, and when applied to the isolation of RNA from *S. salivarius* were found to result in RNA degradation. Mutanolysin is known to have maximal activity at 60°C and is also activated by Mg²⁺ and Ca²⁺ (Yokogawa *et al.*, 1974). Lysis in 10mM potassium phosphate buffer pH6.5 with 10mM MgSO₄ and 1mM CaCl₂ was achieved with 50µg mutanolysin/ml in 15min at 37°C and 5min at 60°C. The reduction in lysis time by the latter procedure allowed undegraded RNA to be isolated. Lysis was also achieved with 50µg mutanolysin/ml when incubating at 37°C in 100mM Tris-HCl buffer (pH 6.8) containing 2mM EDTA. However, as lysis at 60°C in this buffer still required 20min incubation, it was not the buffer of choice.

The procedure that was developed (3.2.1) also utilized DTT as an inhibitor of RNase activity, while selective precipitation by LiCl (Auffray and Rougeon, 1980) was found to be more satisfactory than CsCl gradient centrifugation for purifying the RNA. These two additional modifications resulted in the extraction of *S. salivarius* RNA of a sufficient quality for transcriptional analysis. The removal of proteins by KCl precipitation following cell lysis did not improve the quality of the RNA, but did reduce yields and therefore was not included in the procedure (3.2.1.).

A.



B.

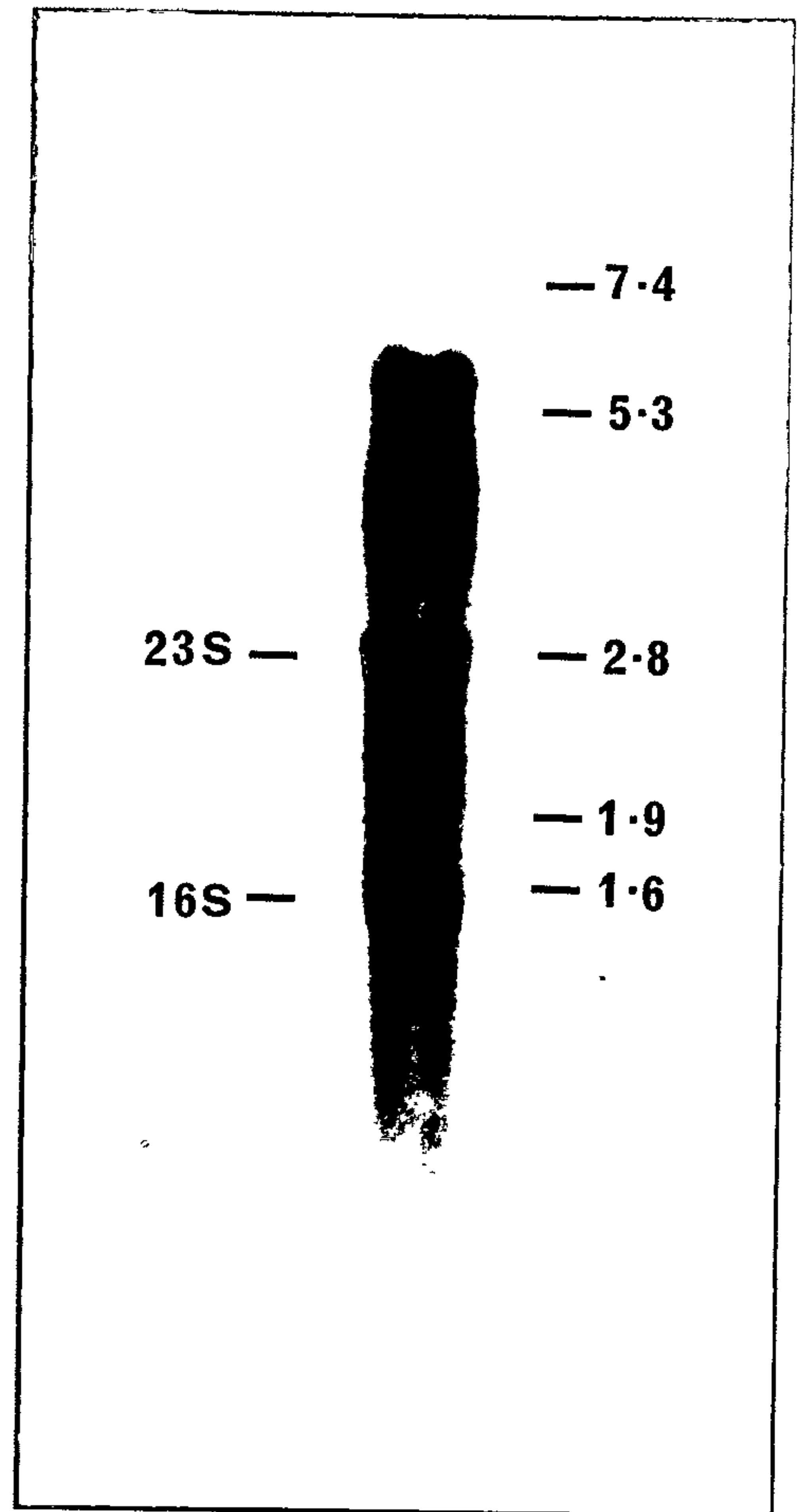


Fig. 3.2. Northern hybridization of total RNA from *S. salivarius* ATCC 25975 using pGSG103 (A) and pGSG203 (B) as probes. The numbers on the right are the molecular mass markers (kb). The positions of the 16S and 23S rRNA bands are also highlighted.

3.3.2 Transcriptional Analysis of the *gtfJK* Locus

Northern blot analysis of the *gtfJK* locus was carried out using internal probes made from pGSG103 and pGSG203 (Fig. 3.1). Both of these probes hybridized to transcripts between 5 and 6kbp in length (Fig. 3.2). The hybridizing transcript for the *gtfK* gene was larger than that of the *gtfJ* gene, corresponding to its longer open reading frame. The size of the transcripts indicated that the genes are transcribed individually rather than as an operon of at least 9.5kbp. No cross-reaction between *gtfJ* and *gtfK* transcripts was observed with either probe. The large size of the transcripts and/or the need to load a high concentration of RNA onto the gel (because of the high percentage of rRNA in the sample) may have had a detrimental effect on the mobility of the mRNA resulting in the measured values of the transcripts being larger than expected. Northern analysis of *gtfJ* gene gave a transcript in the range 4.9-5.5kbp compared with an open reading frame size of 4.6kbp for the gene. Similarly, the range of sizes of the *gtfK* transcript was 5.4-6.0kbp compared with an open reading frame size of 4.8kbp for the gene. RNase protection assays were therefore undertaken to verify the results of the Northern analysis.

Transcription mapping confirmed that the genes were transcribed individually. The probe used in this experiment was made from pGSG202 containing a 1.85kbp *EcoRI* fragment spanning the junction of the two genes (Fig. 3.1). As the two bands obtained in this experiment were both smaller than 1.85kbp (Fig 3.3), it was concluded that the genes were transcribed separately and that there was no detectable readthrough of *gtfK* from *gtfJ*.

3.3.3 Detection of a Putative *gtfJ* Transcriptional Terminator

The sequence between the genes was searched for a potential factor-independent transcription terminator sequence using the 'terminator' application in the GCG package of sequence analysis programs accessed via the ANGIS facility at the University of Sydney. This program analyses sequences according to the algorithm of Brendel and Trivonov (1984). A sequence with a primary score of 4.67, and strong dyad symmetry (secondary structure score of 116) was detected, and is shown in Fig 3.4. This is likely to be the transcription terminator for the *gtfJ*.

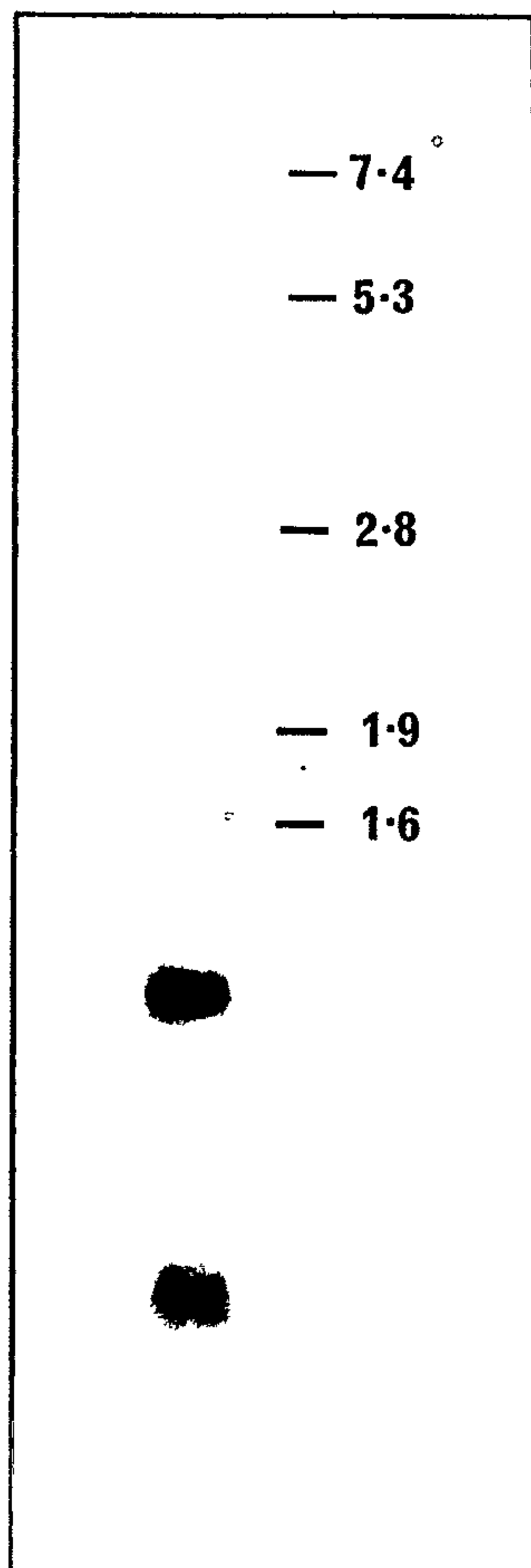


Fig. 3.3. RNase protection assay on total RNA from *S. salivarius* ATCC 25975 using pGSG202 as a probe. The numbers on the right are molecular mass markers (kb).

101 bps downstream from end of *gtfJ*, 109 bps upstream from beginning of *gtfK*.

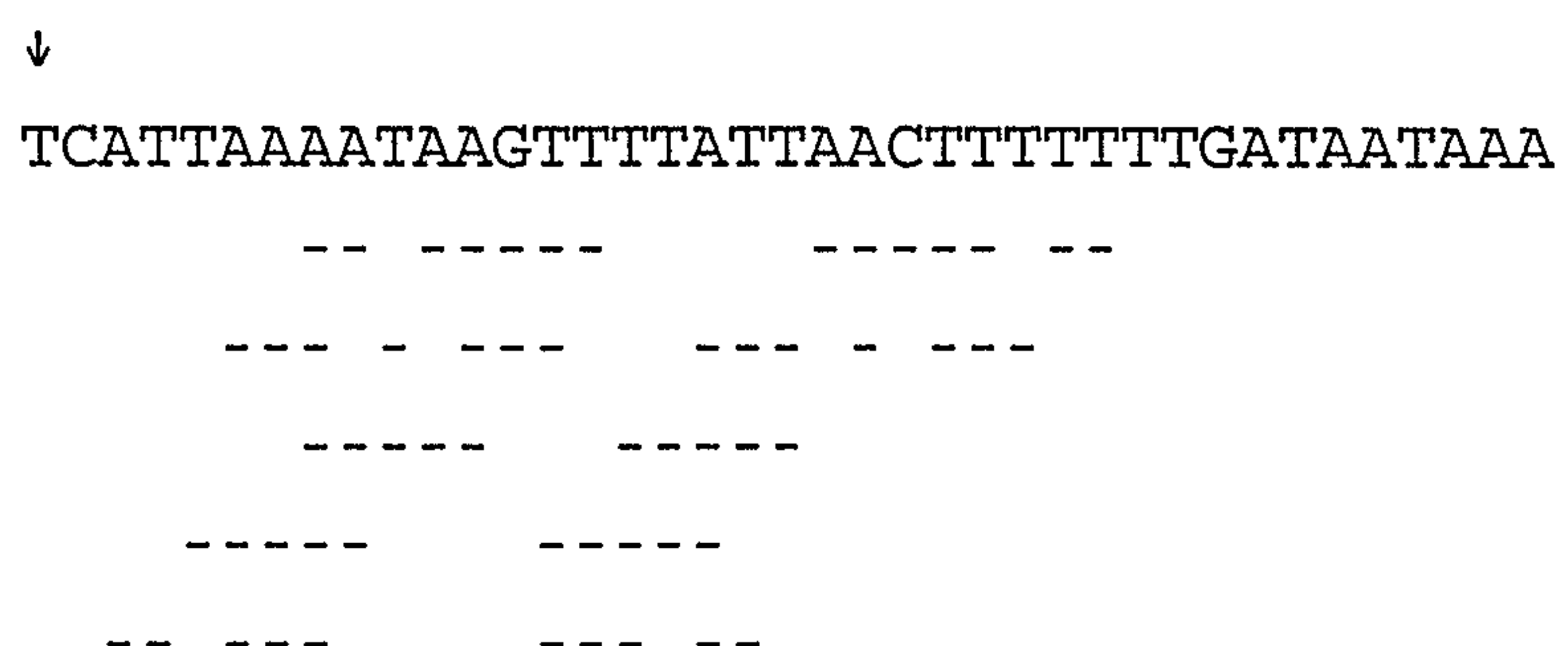


Fig. 3.4. A potential transcript terminator site between the *gtfJ* and *gtfK* genes located by means of the 'Terminator' program (see text). The dashed lines represent regions of dyad symmetry.

3.4 DISCUSSION

There have been only two other reports of transcriptional analysis from oral streptococci, reflecting the difficulty in obtaining high quality RNA from these organisms (Kuramitsu *et al.*, 1991; Shiroza and Kuramitsu, 1993). The procedure described by Shiroza and Kuramitsu (1993) used lysozyme to lyse *S. mutans* and *S. gordonii*, although they did note the RNA could be isolated without lysozyme treatment by boiling the cells in 1% (w/v) SDS (*S. mutans*) or TE buffer (*S. gordonii*). Lysozyme is not effective in the lysis of *S. salivarius* cells. In other streptococci unusual reagents such as phage lysin are sometimes used to affect rapid lysis (Robbins *et al.*, 1987). The method for the preparation of RNA from *S. salivarius* that was developed uses commercially available reagents (mutanolysin is available from Sigma), does not require an ultracentrifugation step and the resultant sample contains no contaminating DNA.

The development of a procedure for the isolation of RNA from *S. salivarius* ATCC 25975 allowed the transcriptional analysis of the *gtfJK* locus. Both Northern hybridizations and RNase protection assay results indicated that the two genes were transcribed separately, with no transcriptional readthrough. A search for a potential factor-independent transcription terminator sequence resulted in the finding of a region of DNA between the *gtfJ* and *gtfK* genes that was likely to be the transcription terminator for the *gtfJ* gene (Fig. 3.4). Recently, nucleotide sequencing has identified a putative promoter elements between the two genes (Giffard *et al.*, 1993).

The *gtfT* and *gtfU* genes of *S. sobrinus* OMZ176 and the *gtfB* and *gtfC* genes of *S. mutans* GS5 are also arranged in tandem (Hanada *et al.*, 1993; Shiroza *et al.*, 1987). In the sequence analysis of the *gtfC* gene a promoter sequence was not identified, leading to the assumption that there was a single promoter for the *gtfBC* locus (Ueda *et al.*, 1988). However, Northern blot analysis has determined that the *gtfB* and *gtfC* genes do not form an operon (Kuramitsu *et al.*, 1991).

Transcriptional activity of the *gtfB* promoter has been studied using chloramphenicol acetyltransferase (CAT) operon fusions (Hudson and Curtiss, 1990; Wexler *et al.*, 1993). The presence of sucrose and variations in growth rate affected GTF expression. GTF activity, however, did not always parallel CAT

activities, leading to the suggestion that post-transcriptional control of enzyme secretion and activity was occurring (Wexler *et al.*, 1993). As the *gtfB* and *gtfC* genes do not form an operon (Kuramitsu *et al.*, 1991), the promoter studied by the use of CAT fusions was that for the *gtfB* gene alone. Since expression of the other two *gtf* genes of *S. mutans*, *gtfC* and *gtfD*, have not been studied, yet another explanation for the results observed with the CAT fusion proteins could be that these other two *gtf* genes are regulated in a different manner by their respective promoters, giving rise to discrepancies between transcriptional activity of the *gtfB* gene and the total GTF activity that was measured in these studies.

Little is known about the regulation of GTFs. Studies by Jacques and coworkers have suggested that phase changes in the membrane lipids may affect GTF production in *S. salivarius* (Jacques, 1985; Jacques *et al.*, 1985; Markevics *et al.*, 1987, Pitty and Jacques, 1987). In these studies, total extracellular GTF activity was determined, and thus any possible effect of various stimuli on expression of individual *gtf* genes, or their respective GTF products remains unknown. This is an important consideration, not only because of the inconsistent results obtained with *S. mutans* (Hudson and Curtiss, 1990; Wexler *et al.*, 1993), but also because of an observation with *S. sobrinus* where Tween 80 has not only been shown to increase the production of GTF-I, but also to promote the production of two new GTFs which synthesize soluble glucan (Walker *et al.*, 1990).

While individual activities are difficult to monitor, the determination of mRNA levels under various growth conditions may in the future give insight into the regulation of these enzymes.

CHAPTER 4

ANALYSIS OF THE REGION UPSTREAM OF THE *gtfJK* LOCUS

4.1 INTRODUCTION

Sequencing downstream of the *gtfJK* locus has shown the presence of another gene (*orf1*) that is transcribed in the same direction as the *gtfJ* and *gtfK* genes (Giffard *et al.*, 1993; Fig 4.1). The presence of this open reading frame (ORF) encoding a hydrophobic protein of molecular weight 12 884 delineates the extent of contiguous *gtf* genes in the downstream region. Since transcriptional analysis of the *gtfJK* locus (Chapter 3) has shown that the *gtfJ* and *gtfK* genes are transcribed separately, and that the size of the *gtfJ* gene transcript is too small to encode for more than one *gtf* gene, it was of interest to investigate the nature of any upstream ORFs.

The 3kbp region immediately upstream of *gtfJ* had been cloned in λ A-8 (Fig. 4.1) and was therefore available for further analysis. This chapter describes the initial analysis of this region and an attempt to isolate an intact *gtf* gene upstream of the *gtfJK* locus.

4.2 METHODS AND MATERIALS

4.2.1 Blotting of Bacteriophage- λ Gene Bank

The gene bank of *S. salivarius* ATCC 25975 previously constructed in λ L47.1 (Pitty *et al.*, 1989) was plated with *E. coli* LE392 in soft agar containing 10mM MgSO₄ on top of LB agar to give approximately 1000 plaques per plate. Plates were incubated overnight (16h) at 37°C. Blotting of plaques onto Hybond N+ nylon membranes (Amersham) followed the manufacturer's protocol. The Hybond N+ membranes were placed on the agar surface for 1min and then removed and placed, colony side up, on a pad of absorbent filter paper soaked in denaturing solution (1.5M NaCl, 0.5M NaOH) for 7min. The membrane was then placed, colony side up, on a pad of absorbent filter paper soaked in neutralizing solution (1.5M NaCl, 0.5M Tris-HCl pH 7.2, 1mM EDTA) for 3min.

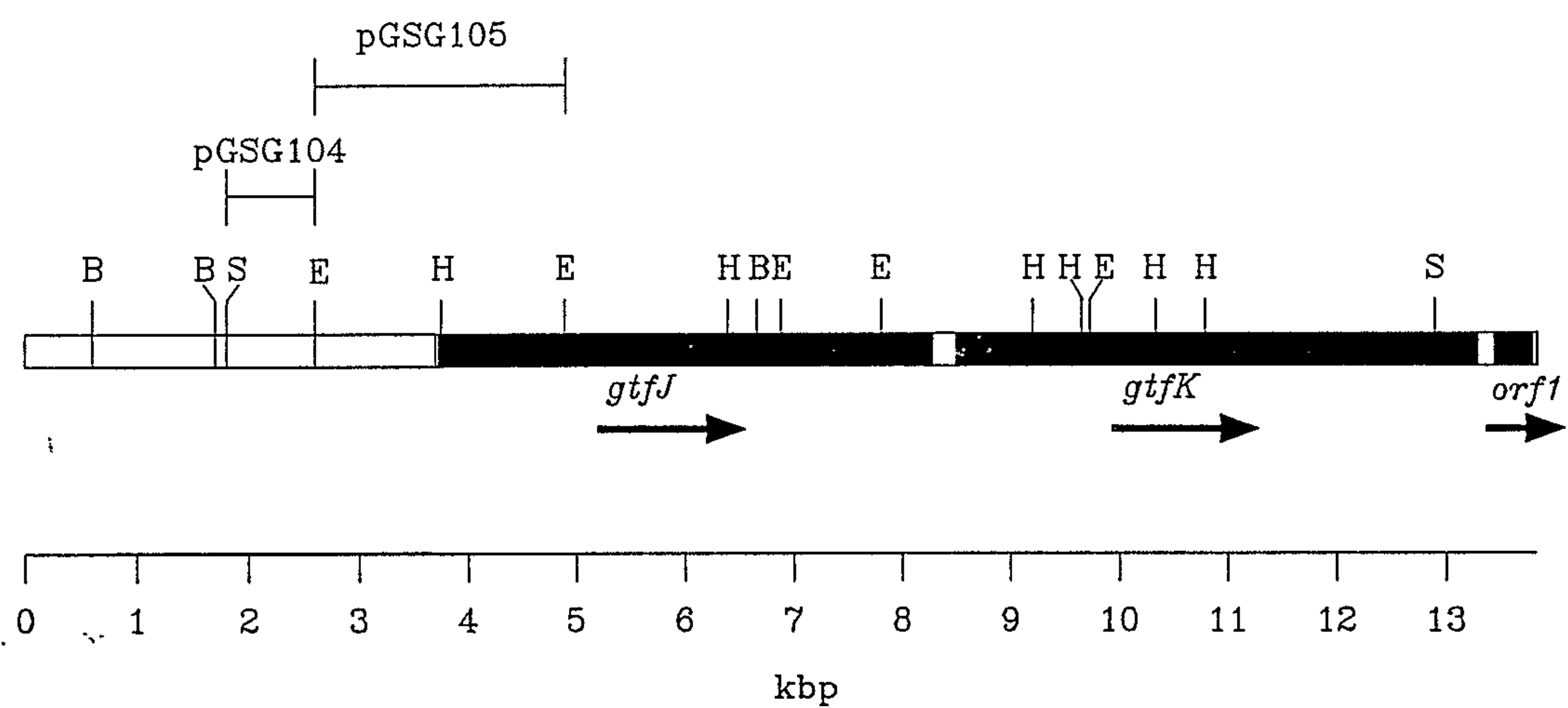


Fig. 4.1. Physical map of the *gtfJK* locus of *S. salivarius* constructed by merging the maps of λ A-8 and λ A-33. The coding regions of *gtfJ*, *gtfK*, and *orf1* are indicated by the filled areas. The DNA sequence of the region upstream of *gtfJ* was not known from previous studies. The restriction fragments from this region cloned into pIBI30 are indicated on the map.

This procedure was repeated with a fresh pad soaked in the same solution. Membranes were then washed in 2 x SSC buffer (2.4.15) and the DNA fixed onto the membrane with 0.4M NaOH (2.4.15) and the membranes washed in 5 x SSC buffer prior to storage in plastic wrap at 4°C until required (0-16h).

4.2.2 Plaque and Southern Hybridizations

Labelling of probes and hybridization of DNA were carried out using one of two methods. In early experiments labelling of probes was accomplished using a Boehringer Mannheim method described in their 'Nonradioactive DNA Labelling and Detection Kit'. This method randomly primes DNA with digoxigenin-dUTP such that after hybridization and washing, hybrid DNA is detected by an enzyme-linked immunoassay that results in the formation of a blue precipitate which can be seen directly on the membrane. Prehybridization (1h) and hybridizations (16h) were carried out at 68°C in 5 x SSC buffer containing 0.5% (w/v) blocking reagent (supplied by Boehringer Mannheim), 0.1% (w/v) N-laurylsarcosine and 0.02% (w/v) SDS. Stringency washes employed 2 x SSC buffer containing 0.1% (w/v) SDS at room temperature (22°C; 2 x 5min washes) followed by 2 x 15min washes at 68°C in 0.1 x SSC buffer containing 0.1% (w/v) SDS. Immunological detection leading to a colour reaction was carried out as described in the kit protocol.

In later experiments and in Southern hybridizations [³²P]-CTP RNA probes were generated using the 'Riboprobe System II' kit (Promega). Hybridizations were carried out as described in the Promega technical manual. Essentially, prehybridizations (1h) and hybridizations (16h) were at 42°C in 0.25M sodium phosphate buffer (pH 7.2) containing 50% (v/v) formamide, 0.25M NaCl, 1mM EDTA, 7% (w/v) SDS and 7% (w/v) PEG. Posthybridization washes were performed in 2 x SSC buffer at room temperature (22°C) for 5min followed by 2 x 20min washes at 65°C in 0.25M sodium phosphate buffer (pH 7.2) containing 2% (w/v) SDS and 1mM EDTA and finally by 2 x 20min washes in 0.04M sodium phosphate buffer (pH 7.2) containing 1% (w/v) SDS and 1mM EDTA. Autoradiographs of the membranes were then obtained. The autoradiographs were aligned with the agar plates and positive plaques removed in an agar plug and eluted in 100µl phage buffer (20mM Tris-HCl pH7.4, 100mM NaCl, 10mM

MgSO₄).

The methods use for the electrophoresis of DNA on agarose gels and transfer of DNA to Hybond N+ membranes for Southern hybridizations have been described previously (2.4.9; 2.4.15).

4.2.3 Production of Bacteriophage- λ Lysates

A culture of *E. coli* LE392 (0.5ml) grown overnight at 37°C in LB medium containing 10mM MgSO₄ was mixed with 50 μ l of eluted phage (4.2.2) and incubated at 37°C for 20min before the addition to 50ml of LB medium containing 10mM MgSO₄. Incubation was continued at 37°C with shaking for 6-7h until the culture had lysed. Chloroform (200 μ l) was added to any cultures that had not lysed in this time and the shaking continued until lysis was complete. These lysates were then used for GTF assays and DNA isolation as described previously (2.4.1, 2.5.1, 2.5.2).

4.2.4 DNA Sequence Analysis

DNA that was to be sequenced was first subcloned into either pIBI30 or pIBI31 (Table 2.1). Sequence determinations were carried out using ssDNA or CsCl purified dsDNA by the chain-termination method of Sanger *et al.* (1977). All sequencing reactions used the Pharmacia T7 sequencing kit according to the manufacturer's instructions using T7 and T3 oligonucleotide primers and [³⁵S]-dATP (Dupont NEN) (see also Section 2.4.18). Sequences were analysed using the IBI Pustell sequence analysis software version 2.03.

4.3 RESULTS AND DISCUSSION

4.3.1 Sequencing Upstream of the *gtfJK* Locus

Sequencing from the *SacI* site of the 0.8kbp *SacI/EcoRI* fragment (pGSG104) within the 3kbp region immediately upstream of *gtfJ* gene revealed homology with a region 3kbp into the *gtfJ* and *gtfK* genes (Fig. 4.2A). This suggested the existence of a third *gtf* gene upstream of the *gtfJK* locus (Fig. 4.2B). As all previously sequenced *gtf* genes are 4.4-4.8kbp in length, it was evident that this

A.

<i>gtfJ</i>	3169	CTAAATACGGTGGTGAATTCTTGGCTGAACTTAAAGCTAAGTACCCT
		* * * * *
pGSG104		C-GACTCTAGAGGATCCCCGGGTACCGAGCTCCAAAAGAAATATCCA
		* * * * *
<i>gtfK</i>	3046	CCAG-TACGGTGGTGCCTTCCTCGACAAACTTCAAAAACCTTTATCCA
<i>gtfJ</i>		GAAATGTTCAAGGTAAACATGATTTCAACTGGTAAACCAATTGATGA
		* * * * *
pGSG104		CAGTTGTTTGAAGATGTGATGATATCAACTGGTAAGAAGATTGATCC
		* * * * *
<i>gtfK</i>		GAAATCTTCAAAGAAGTTATGGAAGCGTCTGGTAAGACAATTGATCC
<i>gtfJ</i>		TTCTGTTAAATTGAAACAATGGAAGGCTGAATACTTCAACGGAACAAA
		* * * * *
pGSG104		ATCTGTTAAAATCAAGCAATGGTCTGCTAAGTACATGAATGGTACTAA
		* * * * *
<i>gtfK</i>		ATCTGTGAAGATTAAACAGTGGGAAGCTAAATACTTTAACGGTACAAA
<i>gtfJ</i>		CGTTCTTGAACGTGGTGGTCTATGTACTTAGCGATG
		* * * * *
pGSG104		TATCCTTGGTCGTGGTAATCGTTATGTTCTCAGTAAC-
		* * * * *
<i>gtfK</i>		CATCCAAAACGTGGTCTGACTATGTTCTCAGCGATG

B.

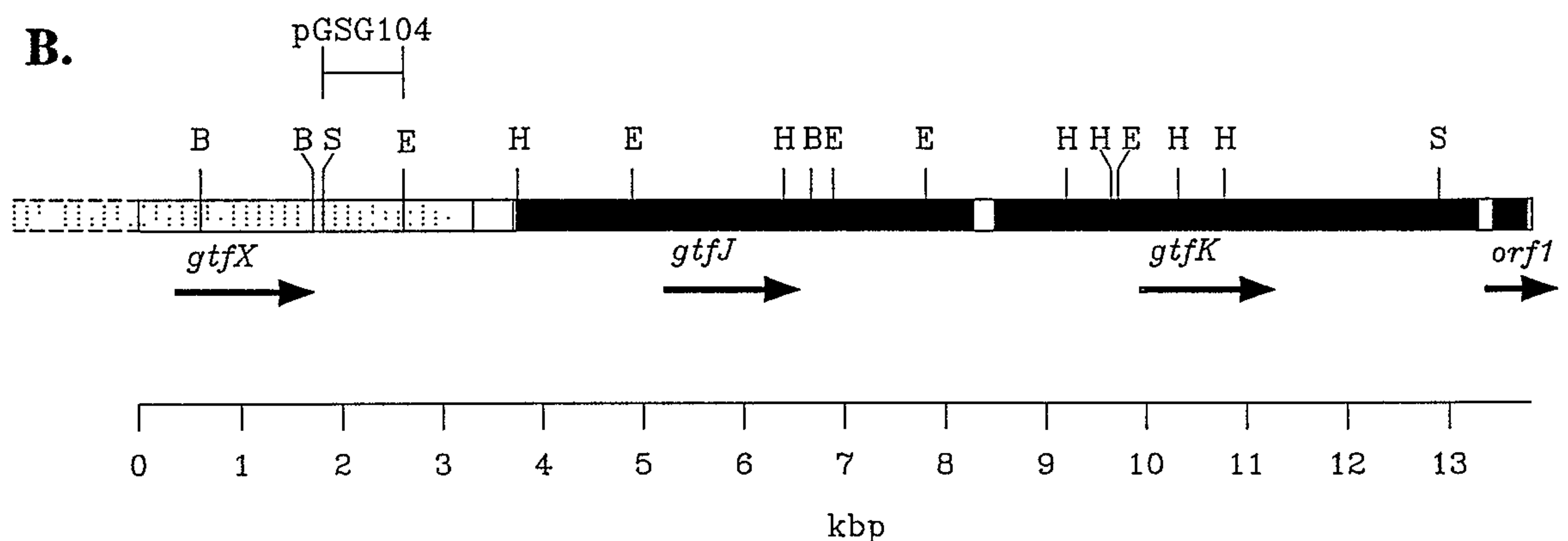


Fig. 4.2. A. Alignment of the sequenced upstream region of the *SacI*/*EcoRI* cloned fragment (pGSG104) with homologous regions of *gtfJ* and *gtfK*.
 B. Position of putative third *gtf* gene (*gtfX*) in relation to the *gtfJ* and *gtfK* genes as determined by sequence alignment assuming similar sizes for all three *gtf* genes. Dashed region indicates part of genome not present on λ A-8.

putative *gtf* gene (designated *gtfX*) could not exist in its entirety in the λ A-8 clone.

4.3.2 Detection of *gtfX* Positive Plaques

The gene bank of *S. salivarius* ATCC 25975 constructed in λ L47.1 (Pitty *et al.*, 1989) was probed with the labelled 0.8kbp *SacI/EcoRI* fragment of λ A-8 in pGSG104 (Fig 4.1) in order to detect any λ -clones containing all or part of this fragment. Use of the digoxigenin labelled probe resulted in a large number of false positives so this technique was abandoned in favour of [³²P]-labelled probes. A number of positive plaques were subsequently isolated and purified by plating out and re-probing. Bacteriophage- λ DNA isolated from these positive clones was digested with *EcoRI* and *SacI* and Southern blot analysis used to confirm the existence of a hybridizing 0.8kbp *SacI/EcoRI* fragment (Fig. 4.3A). Only one hybridizing plaque from the 20 000 thousand probed and the 9 positive plaques isolated contained a fragment of interest that differed in its *SacI/EcoRI* restriction pattern to that of λ A-8. This clone, designated λ B-5.2, exhibited no GTF activity (2.5.2), but gave a positive result in the assay for reducing sugars (2.5.1) whether or not sucrose and/or dextran T10 was present. This latter result raised doubts as to the nature of the insert in λ B-5.2.

The λ B-5.2 clone was also probed with the 2.3kbp *EcoRI* fragment lying downstream of the 0.8kbp *SacI/EcoRI* fragment in λ A-8 (Fig. 4.1). No 2.3kbp fragment was detected in λ B5.2, but rather a 0.9-1.0kbp *EcoRI* fragment of weaker intensity (Fig. 4.3B). It therefore appeared that only a portion of the 2.3kbp *EcoRI* fragment existed in the λ B-5.2 clone, suggesting that the overlap between the inserts of λ A-8 and λ B-5.2 terminated in this region. If the λ B-5.2 clone extended only a short distance (less than 1kbp) into the 2.3kbp *EcoRI* fragment this implied that the putative *gtf* gene did not exist in its entirety in this clone; a fact which was consistent with the lack of GTF activity.

4.3.3 Restriction Mapping of λ B-5.2

The *EcoRI* fragment in λ B-5.2 that hybridized with the 2.3kbp *EcoRI* fragment from λ A-8 was too small to extend into either of the λ -arms as would be necessary if the insert in λ B-5.2 was a true extension upstream of the *gtfJ* gene present in λ A-8. This was the case since the *EcoRI* site in the right arm should have been 2.8kbp downstream of the insert and the equivalent site in the left arm

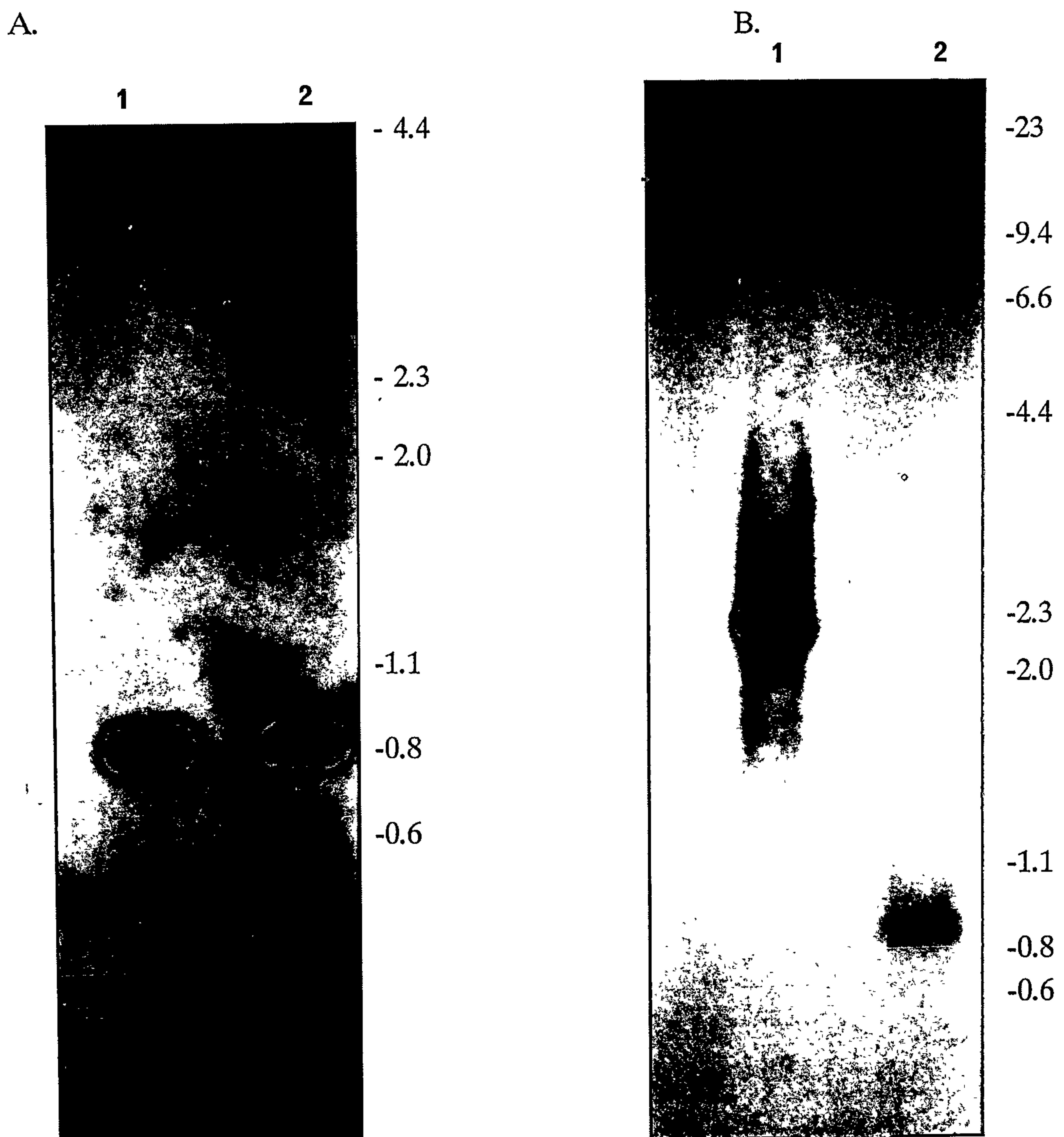


Fig. 4.3. Southern analysis of the λ B-5.2 clone isolated from phage hybridization experiments. The numbers on the right are molecular mass markers (kbp).
A. *SacI/EcoRI* double digested λ B-5.2 (lane 1) and λ A-8 (lane 2) probed with pGSG104 containing the 0.8kbp *SacI/EcoRI* fragment of λ A-8.
B. *EcoRI* digested λ A-8 (lane 1) and λ B-5.2 (lane 2) probed with pGSG105 containing the 2.3kbp *EcoRI* fragment of λ A-8 (cf Fig. 4.1).

1.8kbp upstream of the insert. This meant that any *EcoRI* fragment extending into bacteriophage- λ would need to be at least 1.8kbp in size. However, as deletions in these regions of the λ -arms have previously been noted by us (unpublished observations), this possibility had to be investigated.

4.3.4 Evidence for a Mixed Insert in λ B-5.2

Further restriction mapping of λ B-5.2 suggested that the *Bam*HI site was at the junction of the insert and the right arm of λ L47.1 thus placing the *Sac*I site to the left in the λ B-5.2 insert (Fig. 4.4). Furthermore, when the 6.4kbp *Sac*I/*Bam*HI fragment of λ B-5.2 was cloned into pIBI30 (pGSG306), and CsCl purified dsDNA prepared (2.4.4) and sequenced (2.4.17) from the *Sac*I site, the sequence was found to be identical to that derived from the 0.8kbp *Sac*I/*Eco*RI subclone of λ A-8 (pGSG104) and the 0.8kbp *Sac*I/*Eco*RI (pGSG301; see below) and 1.6kbp *Sac*I/*Hind*III (pGSG302; see Fig. 4.5; see below) subclones of λ B-5.2. This result indicated that the overlap between λ A-8 and λ B-5.2 lay between the *Sac*I and *Bam*HI sites in the central region of the λ B-5.2 insert. This implied that a 'scrambled' clone had been isolated. The restriction map (Fig. 4.4) indicates clearly that the insert in λ B-5.2 contains at least two separate fragments of *S. salivarius* genome, as it is unable to overlap with the λ A-8 insert.

4.3.5 Sequence Similarity Between Regions of the Insert in λ B-5.2 and that in λ A-8

Several restriction fragments from λ B-5.2 were subcloned into pIBI vectors, and ssDNA prepared and sequenced (2.4.5; 2.4.18). Sequencing of two subclones from λ B-5.2, a 0.8kbp *Sac*I/*Eco*RI (pGSG301) and a 1.6kbp *Sac*I/*Hind*III (pGSG302), from the *Sac*I site resulted in the same sequence previously derived from the 0.8kbp *Sac*I/*Eco*RI subclone of λ A-8 confirming the results of Southern blot analysis. These sequences were homologous with DNA that coded for a conserved region of the 'catalytic domain' of GtfJ but showed no homology with the glucan-binding protein of *S. mutans* Ingbritt (Russell *et al.*, 1985). In contrast, sequences obtained from several other subclones could be aligned with the region of the *gtfJ* gene coding for the glucan-binding domain of GtfJ (Fig. 4.5).

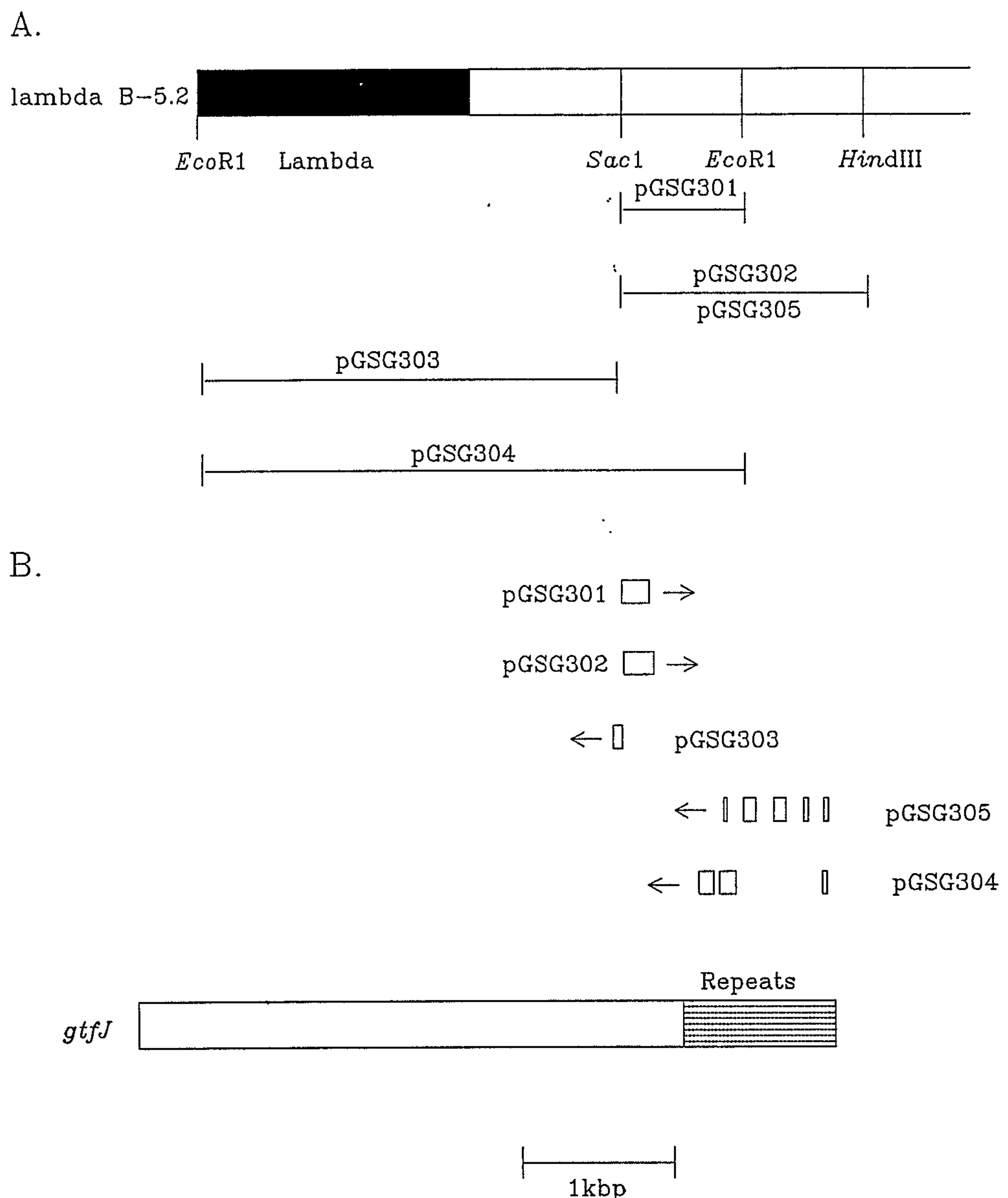


Fig. 4.5. A. Subclones of λ B-5.2 used to determine sequence homology with the *gtfJ* gene. The shaded region represents λ L47.1 DNA.

B. Alignment with *gtfJ* of sequenced subclones of λ B-5.2. Sequences from the *SacI* site of pGSG301 and pGSG302 were identical to the sequence from pGSG104 containing the 0.8kbp *SacI/EcoRI* fragment of λ A-8 (see text). This region showed homology with a region approximately 3kbp into *gtfJ* gene. A sequence from the *SacI* site of the 3.8kbp *SacI/EcoRI* insert in pGSG303 extended this homology upstream. Sequences from the *HindIII* site of pGSG305 and the *EcoRI* site of pGSG304 showed homology with the repeat sequences of the *gtfJ* gene. The arrows indicate the direction of sequencing.

Figs A and B are in the same scale indicating that an entire *gtf* gene could not exist on the λ B-5.2 insert.

4.3.6 Does the *gtfX* gene exist?

The results obtained with the analysis of the λ B-5.2 clone show that it does not encode for an active GTF (4.3.2). Due to the 'scrambled' nature of the insert it was not possible to determine whether or not a complete gene encoding GtfX activity exists upstream of the *gtfJK* locus on the chromosome of *S. salivarius*. Even if the *gtfX* gene does exist, it may, through mutation, be inactive in *S. salivarius*. None of these alternatives, however, were investigated further. In fact, this approach was abandoned due to the inherent difficulty encountered using DNA/DNA hybridization to isolate *gtf* genes as only a single clone from the thousands screened was isolated to which the insert overlapped with that in λ A-8 but which differed from λ A-8. As an alternative, the gene bank of *S. salivarius* ATCC 25975 constructed in λ L47.1 was screened on sucrose-containing media and GTF-positive clones isolated. This work is described in the next chapter.

CHAPTER 5

CLONING OF TWO GENES FROM *STREPTOCOCCUS SALIVARIUS* ENCODING PRIMER-INDEPENDENT GLUCOSYLTRANSFERASE ACTIVITIES

5.1 INTRODUCTION

Early studies on glucan production by oral streptococci suggested that a single GTF was responsible for all glucan synthesis. Guggenheim and Newbrun (1969) looking for an explanation for the variation in polysaccharide production showed that a single strain of *S. mutans* could produce a number of electrophoretically distinct GTFs. Ciardi *et al.* (1976; 1977) were able to show that some of these enzymes synthesize soluble, while others produce insoluble glucans. GTFs are also known to differ in their requirement for primer dextran. There has been difficulty, however, in determining the exact number and nature of GTFs. This has been due to a number of technical problems in the isolation of these enzymes; including the formation of high molecular weight aggregates and proteolytic degradation of these proteins to lower molecular weight but still enzymatically active forms. Cloning of *gtf* genes has allowed the study of individual GTF activities as well as allowing some determination of the numbers and types of GTFs secreted by individual streptococcal species.

Different species of oral streptococci vary in the number and the properties of the GTFs they produce. *S. mutans* has three GTFs, encoded by the genes *gtfB*, *gtfC* and *gtfD* (Pucci *et al.*, 1987; Hanada and Kuramitsu, 1988; Hanada and Kuramitsu, 1989). *S. sobrinus* and *S. downei* each have four GTFs (Takehara, 1984; Russell *et al.*, 1987; Gilmore *et al.*, 1993). In both species two *gtf* genes have been cloned and sequenced (Gilpin *et al.*, 1985; Russell *et al.*, 1985; Russell *et al.*, 1987; Abo *et al.*, 1991; Hanada *et al.*, 1991; Gilmore *et al.*, 1993; Hanada *et al.*, 1993). *S. sanguis* and *S. gordonii* are thought to have single GTFs (Jacques, personal communication), although there have been no reports of the cloning of the *gtf* genes coding for these enzymes.

Isoelectric focusing has revealed that *S. salivarius* ATCC 25975 has at least

three GTFs (Pitty *et al.*, 1989). Recently two genes encoding two different primer-dependent enzymes have been cloned and sequenced (Pitty *et al.*, 1989; Giffard *et al.*, 1991; Giffard *et al.*, 1993). A third primer-independent activity was also known to be secreted by *S. salivarius* as it had been isolated from the culture fluid (Pitty, Simpson and Jacques, unpublished). Although, sequence analysis suggested that a possible third *gtf* gene existed upstream of the *gtfJK* locus, attempts to isolate a clone containing this gene from the gene bank in λ L47.1 by plaque hybridization were unsuccessful (Section 4). An alternative approach, that of screening plaques for primer-independent GTF activities, was therefore used to isolate clones from the bacteriophage- λ gene bank containing *S. salivarius* chromosomal DNA.

5.2 METHODS AND MATERIALS

5.2.1 Screening for Clones Expressing Glucosyltransferase Activity

A gene bank of chromosomal DNA from *S. salivarius* ATCC 25975 previously constructed in λ L47.1 (Pitty *et al.*, 1989) was plated with *Escherichia coli* LE392 in 0.7% (w/v) bacteriological agar containing 10mM MgSO₄ on top of M9 minimal agar medium (a solution of 0.6% [w/v] Na₂PO₄, 0.3% [w/v] KH₂PO₄, 0.05% [w/v] NaCl, 0.1% [w/v] [NH₄]₄SO₄ and 1.5% [w/v] bacteriological agar to which, after autoclaving, was added 0.1mM of sterile CaCl₂ and 2mM sterile MgSO₄; Maniatis *et al.*, 1982) containing 1% (w/v) sucrose, with or without 0.02% (w/v) dextran T-10, and containing those supplements required for the growth of *E. coli* LE392 (0.2% [w/v] glucose and 0.005% [w/v] methionine), to give approximately 1000 plaques/plate. Plates were incubated at 37°C overnight. Plaques suspected of producing GTF activity were detected by opacity. These plaques were removed from the plates in an agar plug and eluted in 100 μ l phage buffer (Section 2.4.1).

5.2.2 Production of Lysates

An aliquot (0.5ml) of a culture of *E. coli* LE392 grown overnight at 37°C in LB medium (Section 2.3) containing 10mM MgSO₄ was mixed with 50 μ l of eluted phage (5.2.1) and incubated at 37°C for 20min before being added to 50ml of

fresh LB medium containing 10mM MgSO₄ at 37°C. Incubation was continued at 37°C with agitation for a further 6-7h until the culture had lysed. Chloroform (200μl) was added to any cultures not lysed in this time and shaking continued until lysis had occurred. The lysates were then used for GTF assays and DNA isolation as described previously (Sections 2.4.1, 2.5.1, 2.5.2).

5.2.4 Digestion of Glucan with Dextranase

Assay mixes (Section 2.5.2; total volume 8ml) containing 3.2ml of λA-8, λA-33, λD-10 or λC-13 lysates, and radiolabelled sucrose were incubated at 37°C for 2h. Dextran T-10 (0.02%; w/v) was included in the assay mix for the two primer-dependent activities, λA-8, encoding GtfJ and λA-33, encoding GtfK. The amount of glucan formed (cpm) by each enzyme was determined by assaying two 500μl aliquots (Section 2.5.2). The remainder of the assay mixes were boiled for 1h to inactivate the enzymes. After cooling to 37°C, *Chaetomium gracile* α-1,6-dextranase (Miles Biotechnology Group) was added to a final concentration of 50mU/ml and the solution incubated at 37°C. Duplicate aliquots (500μl) were removed and assayed for total remaining glucan at varying time intervals over a 5h period. Any reduction in radioactivity (cpm) during this incubation period was attributed to digestion of the glucan by α-1,6-dextranase to smaller products not entrapped under the assay conditions.

5.2.5 Initial Characterization of the Carbohydrate Products Synthesized by the λD-10 Lysate

Insoluble polymer was produced from a λD-10 lysate by adding sucrose (5%, w/v) and thiomersal (0.01%; w/v) to the lysate and incubating at 37°C overnight. The insoluble product was collected by centrifugation (5000g, 10min, 4°C) and washed several times with H₂O before freeze drying. A sample of this product (15mg) was dissolved in 200μl of 1M NaOH and then diluted with H₂O and neutralised with HCl to give a final concentration of 15mg/ml. The total hexose and fructose concentration was then determined using the phenol/sulphuric acid method (Dubois *et al.*, 1956). Essentially 2ml of phenol (2.5%; w/v) and 5ml of 36M sulphuric acid were added to a 10μl sample of the solubilized product. The

mixture was vortexed and allowed to cool to room temperature before measuring the absorbency at 490nm. The A_{490} reading was multiplied by 105 to give the amount of hexose in the sample in μg glucose-equivalents. Fructose was destroyed by boiling a second sample in 2M trifluoroacetic acid for 4h. This sample was assayed, with the difference between the assay results attributable to fructose.

5.2.6 Southern and Northern Hybridizations

S. salivarius chromosomal DNA was available from previous studies (Giffard *et al.*, 1991) and electrophoresis and transfer of DNA and RNA to Hybond N+ membranes was as described in Sections 2.4 and 3.3.

Hybridizations were carried out using different methods dependent upon the gel in question. Southern analysis of the insert encoding *gtfL* used the ECL (Enhanced Chemiluminescence) gene detection system (Amersham). Labelling of *Bam*HI cut pGSG403 and detection on Hyperfilm-ECL were carried out using reagents supplied by Amersham according to the manufacturer's instructions. Prehybridization (4h) and hybridization (16h) were carried out at 42°C in Amersham hybridization buffer containing 5% (w/v) Amersham blocking agent and 0.5M NaCl. Post-hybridization washes were 2 x 20min washes at 42°C in 0.5 x SSC buffer (Section 2.4.15) containing 6M urea and 0.4% (w/v) SDS followed by a 5min wash at 22°C in 2 x SSC buffer.

For the Northern analysis of the *gtfL* transcript, the prehybridization (4h) and hybridization (16h) were carried out at 68°C in 0.5M sodium phosphate buffer (pH 7.2) containing 1mM EDTA and 7% (w/v) SDS as described by Church and Gilbert (1984) to which was added 10% (w/v) dextran sulphate. The stringency washes were 2 x SSC buffer containing 0.1% (w/v) SDS at room temperature (10min), followed by 1 x SSC buffer containing 0.1% (w/v) SDS at 68°C for 15min and a final wash with 0.1 x SSC buffer containing 0.1% (w/v) SDS at 68°C for 20min. A probe made from pGSG403 was labelled with [α - ^{32}P]dCTP (Dupont NEN) using the Promega 'Prime-a-Gene' kit according to the manufacturer's instructions.

Southern and Northern-analysis of *gtfM* were carried out using [α - ^{32}P]CTP

(Dupont NEN) labelled riboprobes. Prehybridizations (4h) and hybridizations (16 h) were carried out at 42°C in 0.25mM sodium phosphate buffer (pH 7.2) containing 50% (v/v) formamide, 1mM EDTA, 7% (w/v) SDS and 7% (w/v) PEG. The subclone pGSG501 linearised with *Xho*1 was labelled with [α -³²P]CTP using T3 RNA polymerase (Promega) and the Promega 'Riboprobe System II' kit according to the manufacturer's instructions (see Section 3.2.4). Posthybridization washes were in 2 x SSC buffer at room temperature for 5min, 2 x 20min washes at 65°C in 0.25M sodium phosphate buffer (pH 7.2) containing 2% (w/v) SDS and 1mM EDTA followed by 2 x 20min washes at 65°C in 40mM sodium phosphate buffer (pH 7.2) containing 1% (w/v) SDS and 1mM EDTA.

5.2.7 SDS-PAGE and Isoelectric Focusing

E. coli NM522 transformed with phagemids containing one of the four *gtf* genes (pGSG101, pGSG201, pGSG401, pGSG501) were grown overnight in a shaking incubator at 37°C in 200ml aliquots of LB medium supplemented with ampicillin (100µg/ml). Control cultures of the host strain, *E. coli* NM522, were also grown in 200ml aliquots of LB medium under the same conditions. Cells obtained by centrifugation (5000g, 15min, 4°C), were resuspended in 50ml of 20mM potassium phosphate buffer (pH 6.5). Cells were disrupted by sonication (Branson Sonic Power) at 40W using 60sec pulses, with cooling on ice between pulses, for a total sonication time of 5min.

S. salivarius ATCC 25975 was grown at 37°C to mid-exponential phase in 200ml of SDM containing 0.05% (v/v) Tween 80. Cells were harvested by centrifugation (5000g, 15min, 4°C) and the supernatant concentrated 10 fold in a hollow fibre concentrator (Amicon).

Samples (100µl) of sonicates or concentrated culture fluid were analysed by SDS-PAGE as previously described (Section 2.5.3). Following electrophoresis, the gels were washed in 50mM Tris-HCl (pH 7.5) and incubated in a sucrose solution (Section 2.5.4) containing 0.02% (w/v) dextran T-10 for 2h before PAS staining as previously described (Section 2.5.4).

Isoelectric focusing was carried out using Ampholine PAGplate (Pharmacia) utilizing the Hoefer Isobox System (Hoefer). Electrode solutions

were 1M H₃PO₄ and 1M NaOH. The PAGplates were cut in half, placed on the Isobox and electrode strips applied. Samples (50-100µl) and markers (20µl) (Pharmacia Isoelectric Focusing Calibration Kit pH 3-10) were applied onto sample applicators close to the positive electrode. The Isobox temperature was set at 5-7°C and power applied to the gel at 4W for 30min, followed by 8W for 30min and finally 12W for 30min. Marker lanes were removed and fixed in 11.5% (w/v) trichloroacetic acid containing 3.5% (w/v) sulphosalicylic acid prior to the gel being stained as described in the Ampholine PAGplate manual (Pharmacia). The remaining gel (containing the GTFs) was washed in 50mM Tris-HCl (pH 7.5) for 90min and incubated overnight in a sucrose solution (Section 2.5.4) containing 0.02% (w/v) dextran T-10. Polymer forming activities were then detected using the PAS stain (see Section 2.5.4).

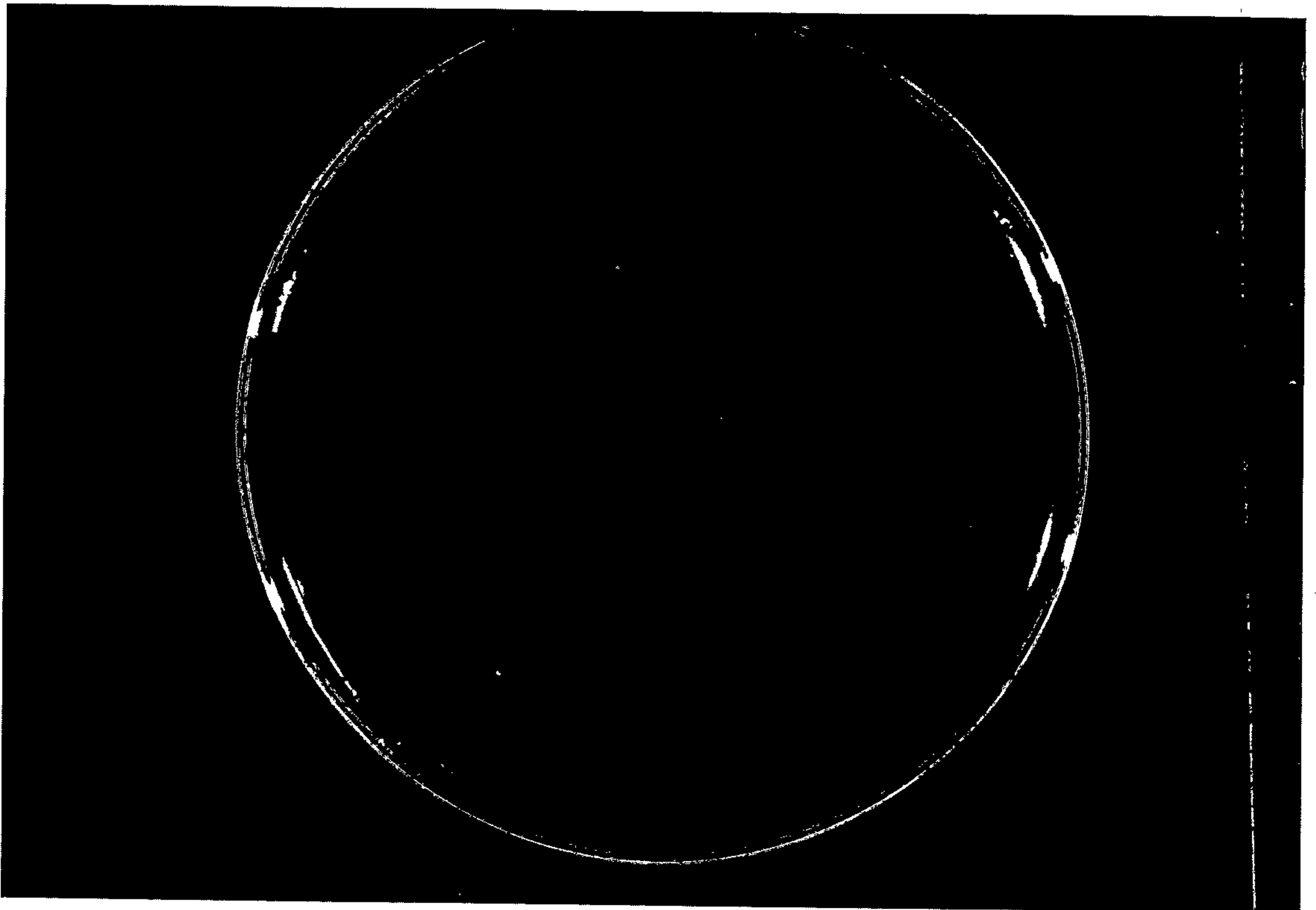
5.3 RESULTS

5.3.1 Detection of Polymer-Forming Bacteriophage-λ on Sucrose Agar Plates

Seventy-two plaques were picked off minimal media plates containing both sucrose and dextran T-10. The plaques showed a variety of morphologies including clear plaques with a cloudy halo, cloudy plaques with or without a halo and raised opaque plaques. DNA was isolated from a representative 32 recombinants, showing varying morphologies, and digested with *Eco*RI. Following electrophoresis, restriction patterns were compared with the bacteriophage-λ recombinants, λA-8 and λA-33, containing the *gtfJ* and *gtfK* genes respectively. The majority of the recombinants had *Eco*RI restriction patterns identical to that of the λA-8 coding for GtfJ activity. However, two other restriction patterns were noted. λD-10 possessed a raised opaque plaque morphology (Fig. 5.1A), whereas λD-40 appeared as a clear plaque with a heavy halo. λD-40 exhibited no glucosyltransferase activity, while λD-10 appeared to code for a primer-independent GTF activity that produced an insoluble product (Table 5.1).

GTF assays were also carried out using lysates dialysed in 10mM sodium citrate buffer (pH 6.0) and by replacing potassium phosphate buffer in the assay mix with the same sodium citrate buffer. Since the enzyme was active under these

A.



B.

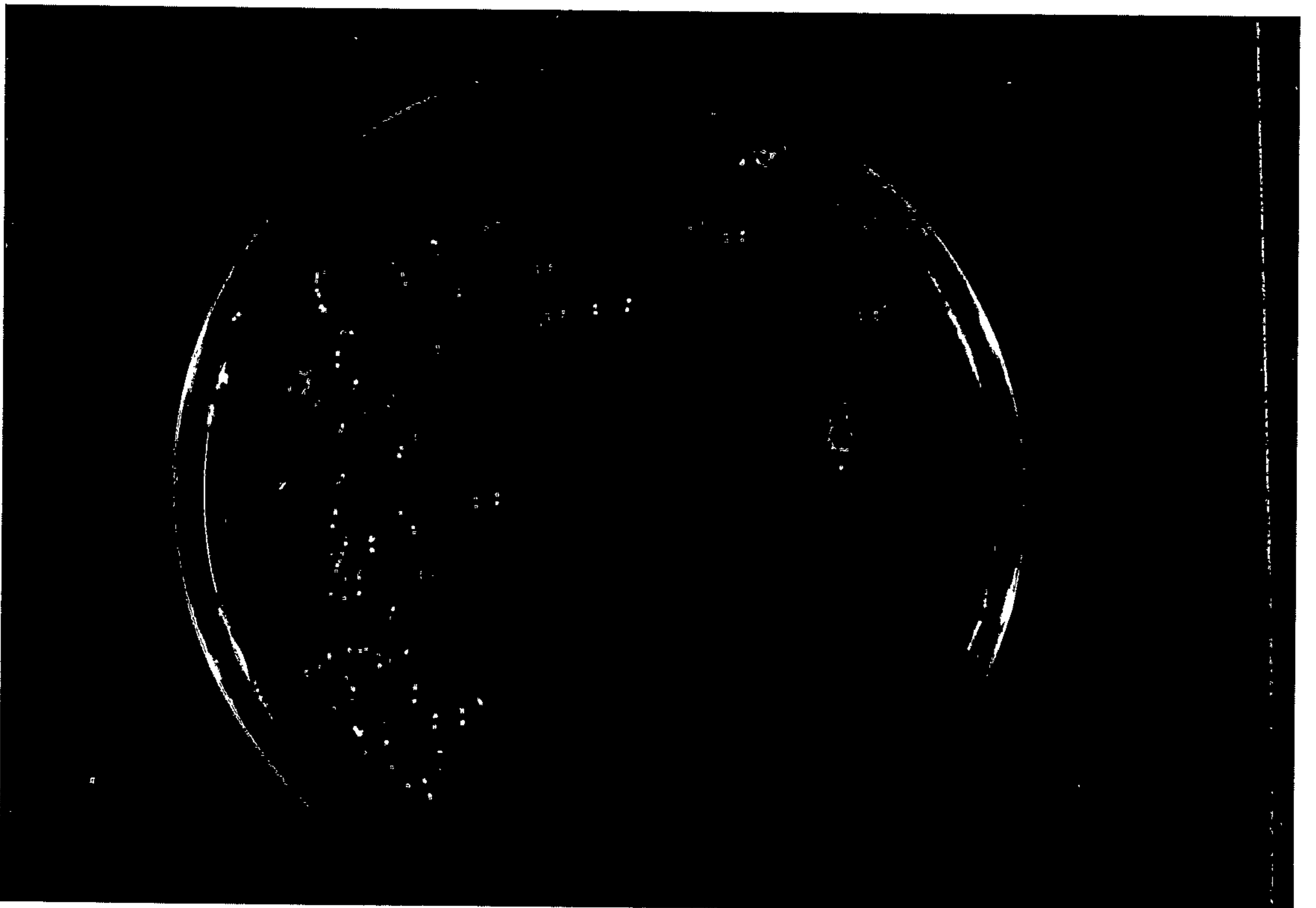


Fig. 5.1. Morphology of bacteriophage- λ recombinants, λ D-10 (A) and λ C-13 (B), grown on an *E. coli* LE392 lawn on M9 minimal salts agar containing sucrose.

Table 5.1. Glucan Production by Bacteriophage- λ Recombinants

Bacteriophage- λ Recombinant	Glucan Production (cpm)			
	-Dextran T-10 Insoluble	-Dextran T-10 Soluble	+Dextran T-10* Insoluble	+Dextran T-10 Soluble
λ A-8	0	0	2251	0
λ A-33	0	322	0	820
λ D-10	1735	0	602	0
λ D-40	0	0	0	0
λ C-13	0	554	0	638

* the concentration of dextran T-10 was 0.02% (w/v)

Table 5.2. Effect of Assay Buffer on Glucan production (cpm) by λ D-10

Buffer	Glucan Production (cpm)	
	-Dextran T-10	+Dextran T-10*
Potassium Phosphate	1490	418
Sodium Citrate	1617	454

* the concentration of dextran T-10 was 0.02% (w/v)

phosphate-free conditions (Table 5.2) this confirmed that it was not a sucrose phosphorylase, which produces an insoluble product in the presence of phosphate and has led to confusion in defining streptococcal cloned gene products in the past (Russell *et al.*, 1988). The GTF encoded by λ D-10 was designated GtfL.

Twenty plaques were also picked off minimal media plates containing sucrose and no dextran T-10. Plaque morphologies included opaque raised, clear with a halo and one plaque with raised mucoid polymer. All recombinants had sucrase activities as evident from the microtitre reducing sugar assays (Section 2.5.1). DNA was isolated from all 20 recombinants and *Eco*R1 restriction patterns compared with the previous isolates, λ A-8, λ A-33, λ D-10 and λ D-40. Several recombinants had the same restriction pattern as λ D-10, and a single isolate was identical to λ A-8. Of those recombinants with different restriction patterns only λ C-13 exhibited GTF activity. This recombinant produced a mucoid polymer on sucrose plates (Fig. 5.1B). This clone coded for a primer-independent GTF activity that produced a soluble product (Table 5.1). This GTF activity encoded by λ C-13 was designated GtfM.

5.3.2 Preliminary Analysis of Glucans Formed by the Various GTF Activities of *S. salivarius*

The glucans produced by GtfK (λ A-33) and GtfM (λ C-13) were readily degraded by the α -1,6-dextranase of *Chaetomium gracile*, whereas the GtfJ (λ A-8) and GtfL (λ D-10) products were not (Table 5.3). The results for the digestion of the glucans produced by GtfK and GtfJ were consistent with a previous study (Pitty *et al.*, 1989).

The product of GtfL was extremely insoluble and refractory to subsequent simple quantitative analysis. Attempts to dissolve the product in 0.5M H₂SO₄ were unsuccessful, even with boiling for 4h. The product was soluble in 1M NaOH, though neutralisation or dilution with H₂O led to precipitation. Despite these problems it was possible to show qualitatively that the product was a glucan. Analysis of a 150 μ g sample suggested that 58.2% was hexose. Trifluoroacetic acid treatment reduced this value to 45.4%. As some of the product precipitated out of solution during the latter treatment, this might imply that the product contained

Table 5.3. *Chaetomium gracile* α -1,6-dextranase digestion of glucans produced by the GTFs of *S. salivarius*.

Time (h)	Digestion of Glucan Produced By:			
	GtfJ	GtfK	GtfL	GtfM
0	3769*	4098	1744	2877
1	3670	3056	1790	119
2	4025	1081	1816	57
5	3671	732	1756	92

* Amount of glucan (cpm) remaining after a given period of incubation in the presence of the α -1,6-dextranase (for further details, see text).

negligible, if any, fructose. Further HPLC analysis of the untreated product by Dr N. Packer at the School of Chemistry, Macquarie University, confirmed that the only sugar in the product was glucose (see also Chapter 7 on ^{13}C -NMR analysis).

5.3.3 Endonuclease Restriction Maps of the Chromosomal DNA in $\lambda\text{D-10}$ and $\lambda\text{C-13}$

Endonuclease restriction maps of $\lambda\text{D-10}$ coding for GtfL and $\lambda\text{C-13}$ coding for GtfM were constructed using double restriction digests (Fig. 5.2). These maps showed that the *gtfL* gene was located on an 11kbp insert and the *gtfM* gene on an 8.3kbp insert. Neither of these clones contained a *SacI* site indicating that they were not candidates for the putative *gtfX* clone, which contains a 0.8kbp *SacI/EcoRI* fragment (see Chapter 4).

5.3.4 Subcloning of the *gtf* Genes

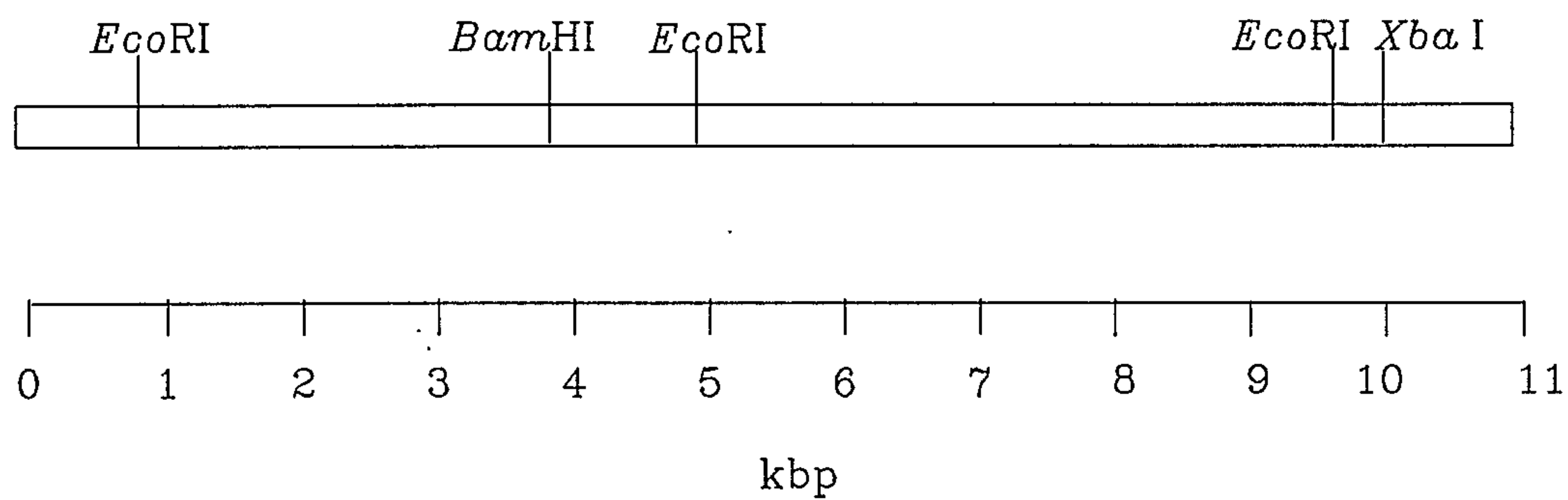
Previously sequenced *gtf* genes had a coding region of about 4.5kbp. Restriction fragments greater than 4kbp were subcloned into the phagemid vectors pIBI30 and pIBI31. The 6.2kbp *Bam*HI/*Xba*I fragment from $\lambda\text{D-10}$ was cloned into both phagemid vectors to form pGSG401 and pGSG402 (Table 2.1). GTF activity was only expressed in the pIBI30 subclone (i.e. pGSG401). This would suggest that the promoter for the *gtfL* gene did not lie within this fragment and thus GtfL expression was dependent upon the phagemid promoter for gene expression. As the pIBI30 promoter lies upstream of the *Bam*HI site this indicates that the gene is transcribed towards the *Xba*I site. The subclones containing the 4.8kbp and 4.1kbp *Eco*R1 fragments from $\lambda\text{D-10}$ in pIBI30 (pGSG403 and pGSG404 respectively; Table 2.1) demonstrated no sucrase or polymer forming activity.

The 5.3kbp *Xba*I fragment from $\lambda\text{C-13}$ subcloned into pIBI31 (pGSG501) was positive for GTF activity as was the 4.8kbp *Xba*I/*Eco*RI from $\lambda\text{C-13}$ subcloned into pIBI31 (pGSG502). The 3.7kbp *Kpn*I/*Xba*I fragment of $\lambda\text{C-13}$ subcloned into pIBI31 (pGSG503) did not demonstrate sucrase or polymer forming activity.

5.3.5 Southern and Northern Hybridization Analysis

Southern and Northern hybridizations were carried out to ensure that there were

A.



B.

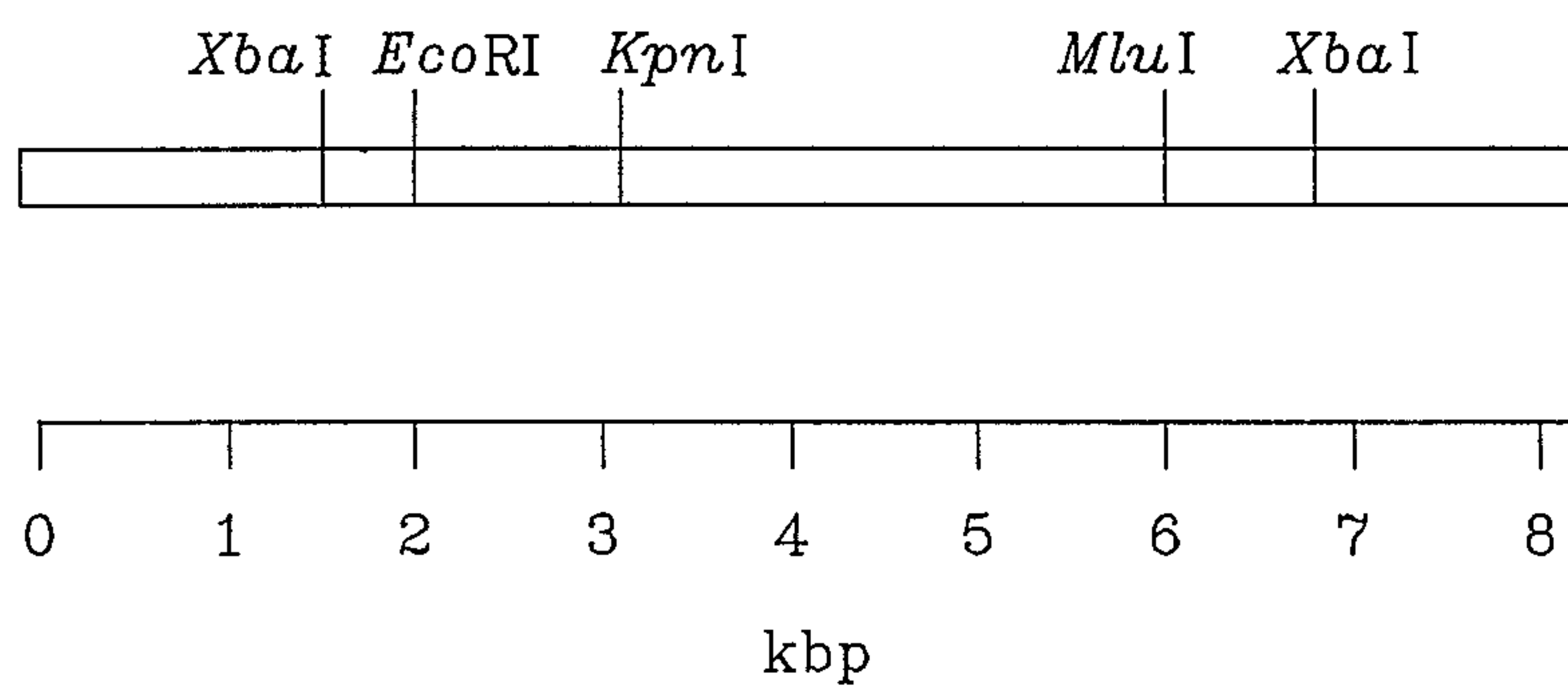


Fig. 5.2. Restriction maps of the inserts in λ D-10 (A) and λ C-13 (B) constructed using double restriction digests. No *Sac*I sites were found.

no cloning artefacts and that the *gtf* genes encoded by the DNA inserts in the phagemid vectors were expressed by *S. salivarius*.

Both genes were expressed in *S. salivarius* as shown by Northern blot analysis (Fig. 5.3). Total RNA was probed with the pGSG403 containing the 4.8kbp *Eco*RI fragment of λ D-10 and pGSG501 containing the 5.3kbp *Xba*I fragment of λ C-13. Both of these probes hybridized to the 16S and 23S ribosomal bands and to a larger transcript. The *gtfL* transcript (Fig. 5.3A) was between 5 and 6kbp in length. Unfortunately the result for *gtfM* was not as clear. Despite numerous attempts, hybridizing bands always showed a lower intensity than similar hybridizing bands obtained in Northern analysis of the *gtfL* and *gtfJ* and *gtfK* genes (Section 3.3.1). Unfortunately, these bands show even more poorly in the photograph shown in Fig. 5.384. The detection of ribosomal bands, presumably through attachment to mRNA, confirmed, however, that the *gtfM* gene was expressed in *S. salivarius*.

Southern analysis confirmed the origin of both inserts (Fig. 5.4). The pGSG403 subclone containing the 4.8kbp *Eco*RI fragment of λ D-10 was used to probe *Eco*RI digested *S. salivarius* chromosomal DNA and *Eco*RI digested pGSG401 containing the 6.2kbp *Bam*HI/*Xba*I fragment of λ D-10. A single hybridizing fragment of 4.8kbp was detected from these digests, indicating that the 4.8kbp *Eco*RI fragment existed in the *S. salivarius* chromosome.

The pGSG501 subclone containing the 5.3kbp *Xba*I fragment of λ C-13 was used to probe *Xba*I/*Mlu*I digested pGSG501 and *S. salivarius* chromosomal DNA. A strongly hybridizing fragment of 4.5kbp were detected from these digests, indicating that this region of the λ C-13 insert existed in *S. salivarius* chromosomal DNA (Fig. 5.5).

5.3.6 SDS-PAGE and Isoelectric Focusing Analysis

SDS-PAGE analysis of the phagemids expressing GTF activity indicates that all four *gtf* genes coded for proteins which produce products in the presence of sucrose and dextran T-10 that stained with the PAS stain (Fig. 5.6). Each of the GTFs expressed by the phagemids appeared to have undergone proteolysis with a number of bands staining with the PAS stain. The molecular weight of the

A.

B.

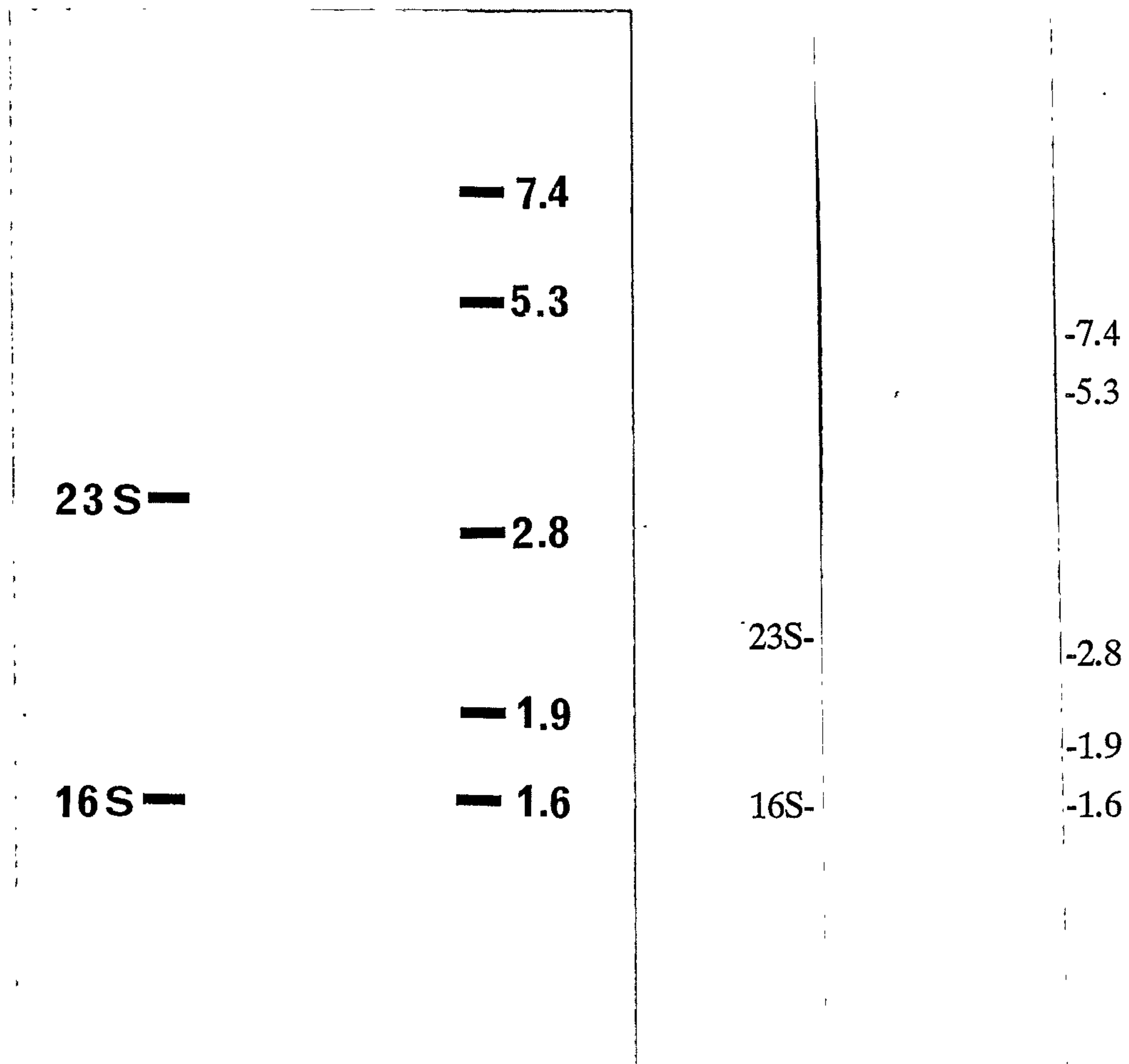


Fig. 5.3. Northern hybridization of total RNA from *S. salivarius* ATCC 25975 using pGSG403 containing the 4.8kbp *Eco*RI fragment from λ D-10 (A) and pGSG501 containing the 5.3kbp *Xba*I fragment from λ C-13 as probes (B). The numbers on the right are the molecular mass markers (kb). The positions of the 16S and 23S rRNA bands are also highlighted.

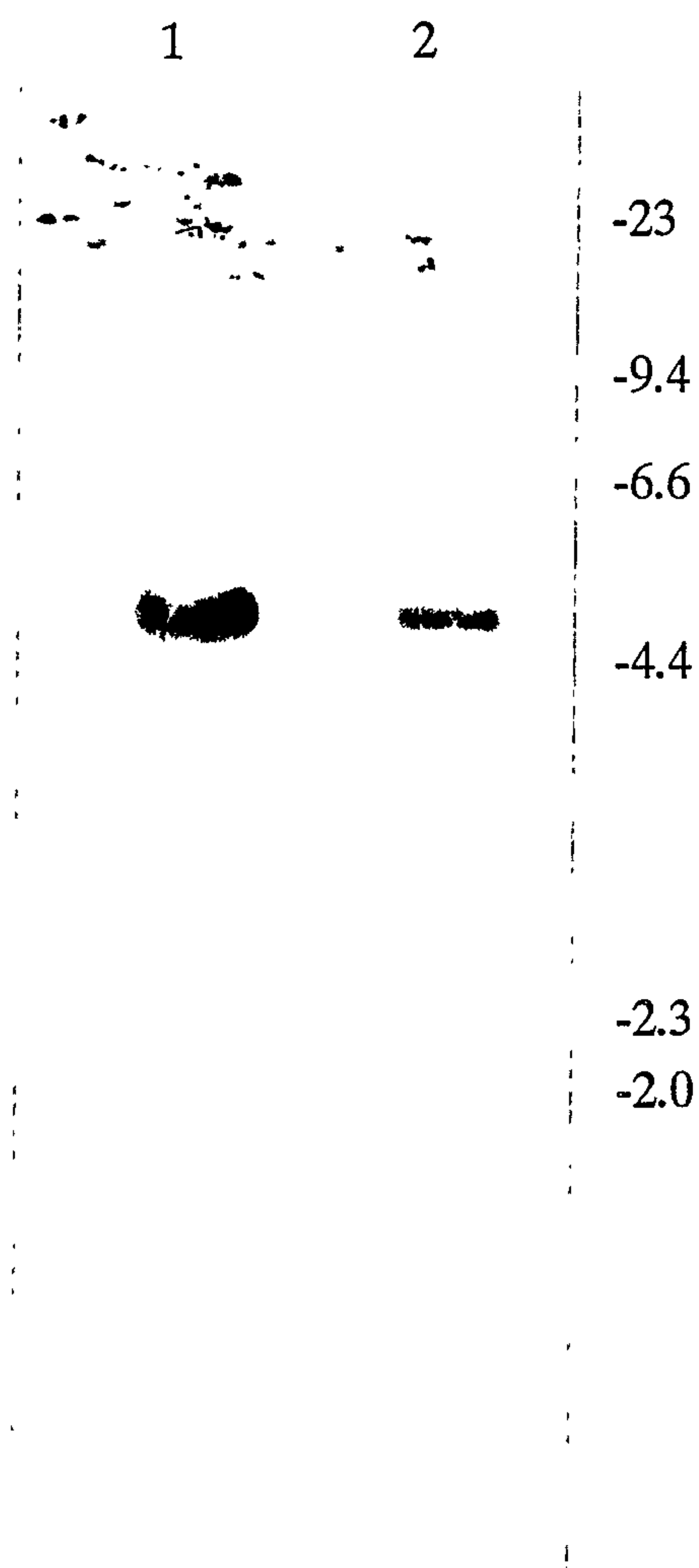


Fig. 5.4. Southern hybridization analysis of the *S. salivarius* chromosome using pGSG403 containing the 4.8kbp *Eco*R1 fragment of λ D-10 to probe the *Eco*R1 digests of pGSG401 containing the 6.2kbp *Bam*H1/*Xba*I fragment of λ D-10 (lane 1), and *S. salivarius* chromosomal DNA (lane 2). The numbers on the left are the molecular mass markers (kbp).

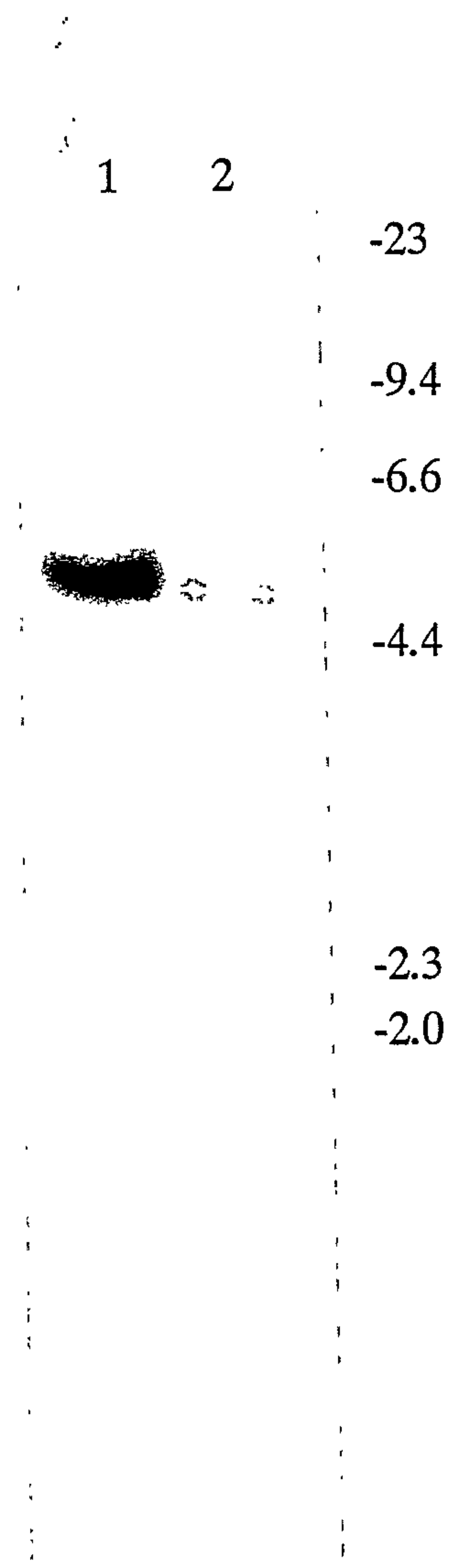


Fig. 5.5. Southern hybridization analysis using pGSG501 containing the 5.3kbp *Xba*I fragment of λ C-13 to probe the *Xba*I/*Mlu*I digested pGSG501 (lane 1) and *S. salivarius* chromosomal DNA (lane 2). The numbers on the left are the molecular mass markers (kbp).

major band(s) for each recombinant as well as the range of molecular weights detected by the PAS staining procedure were determined (Table 5.4). The GtfM in particular appeared to be easily degraded prior to or during SDS-PAGE, as there were a large number of bands that stained with the PAS stain. This result was consistently obtained using a number of fresh sonicates. The bands stained an intense purple similar to those of the GtfK bands, suggesting the products possessed similar structure. Most of the GtfL activity was in a single broad band similar in colour to the bands of GTF activity secreted by *S. salivarius* and the bands produced by GtfJ. The molecular weight of these bands, however, were smaller than the GTFs expressed by the other cloned *gtf* genes and the activities present in the culture fluid of *S. salivarius*. The GtfJ activity is thought to produce glucan of essentially α -1,3-linked glucosyl residues which do not give a colour reaction in PAS staining. However, dextran T-10 is required for this activity and thus the product contains some α -1,6-linked glucosyl residues.

Isoelectric focusing indicated that the cloned GtfJ, GtfK and GtfL activities all had acidic pIs of approximately 5.1-5.4, slightly higher than the major band of activity (pI 5.3) found in the culture fluid of *S. salivarius* (Fig. 5.7, Table 5.4). Similar pI values for GtfJ and GtfK have been obtained previously (Pitty *et al.*, 1989). The pI(s) of GtfM were difficult to obtain. Despite a poor photographic record (Fig. 5.7, lane 5) GtfM was shown to possess an activity with a pI of 5.4. Another activity with a pI of 9.2 was also apparent (Table 5.4). However, there was no corresponding band in the *S. salivarius* culture fluid suggesting it may have been a proteolytic digestion product.

5.4 DISCUSSION

E. coli does not utilize sucrose and is unable to transport it into the cell. Previous studies have indicated that the C-terminus of cloned GTFs prevents their secretion into the periplasm of *E. coli* (Hanada *et al.*, 1991; Kato and Kuramitsu, 1991). Thus detection of cloned sucrase or GTF activities would normally require lysis of the *E. coli* cells containing plasmids expressing GTF activity. However, this problem is overcome when using bacteriophage- λ recombinant clones. In this instance, the release of glucose and fructose from plaques hydrolysing sucrose

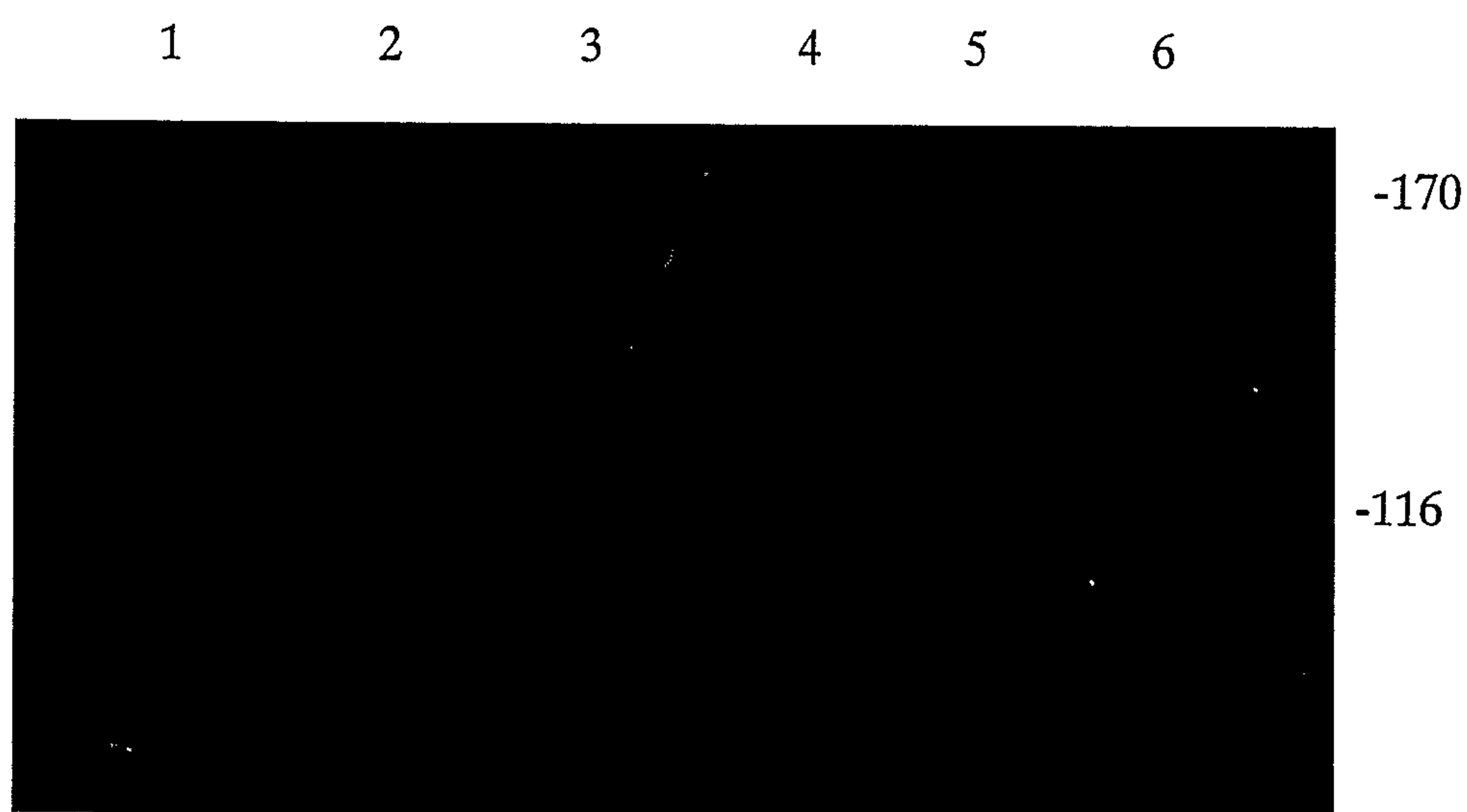


Fig. 5.6. Periodic acid-Schiff staining for GTF activity following SDS-PAGE analysis. Culture fluid of *S. salivarius* grown in the presence of 0.05% (v/v) Tween 80 (lane 1); *E. coli* NM522 with pGSG101 expressing GtfJ (lane 2); *E. coli* NM522 with pGSG201 expressing GtfK (lane 3); NM522 with pGSG401 expressing GtfL (lane 4); NM522 with pGSG501 expressing GtfM (lane 5); *E. coli* NM522 devoid of plasmid (lane 6). The number on the right are molecular weight markers.

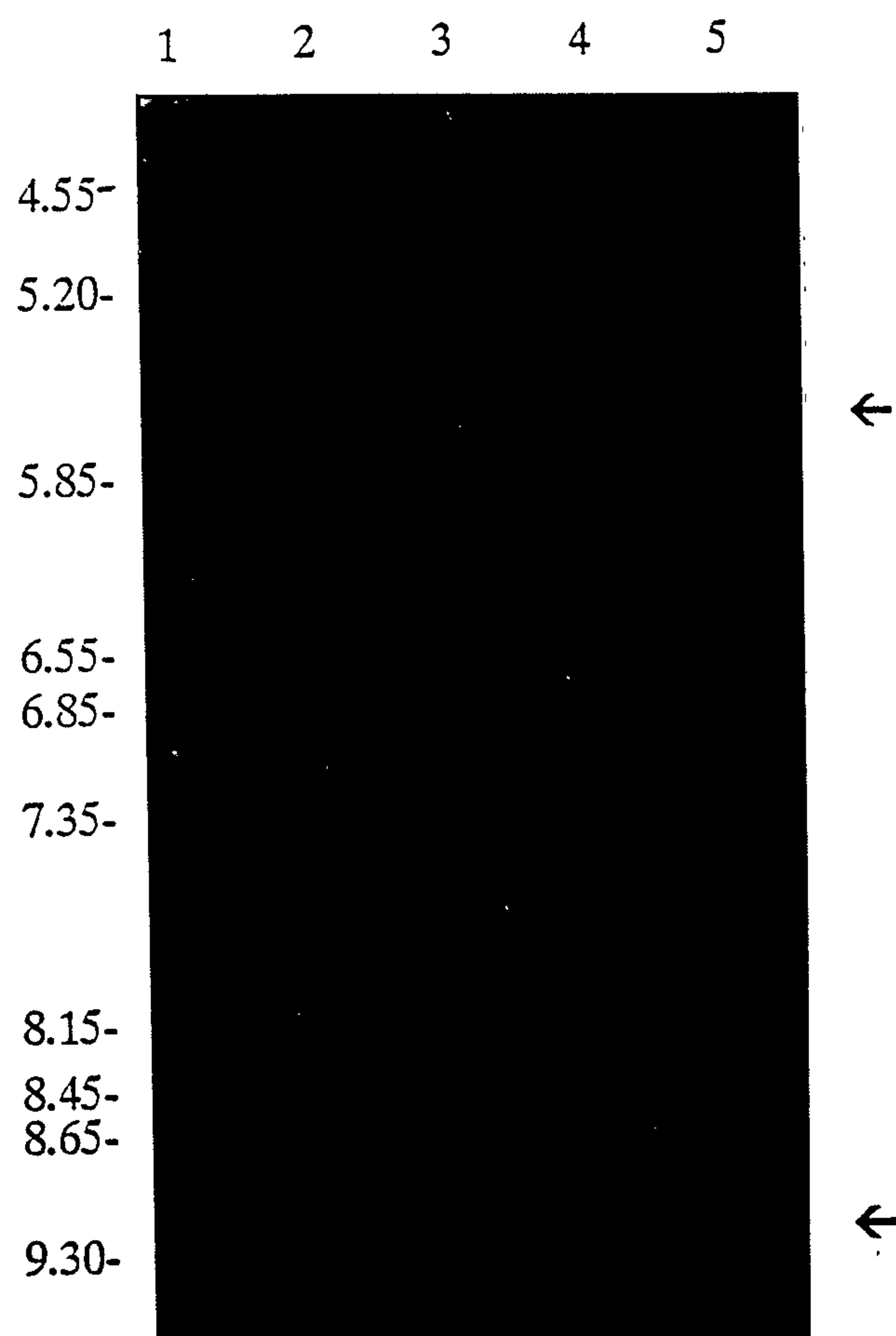


Fig. 5.7. : Periodic acid-Schiff staining for GTF activity following isoelectric focusing. Culture fluid of *S. salivarius* grown in the presence of 0.05% (v/v) Tween 80 (lane 1); *E. coli* NM522 with pGSG101 expressing GtfJ (lane 2); *E. coli* NM522 with pGSG201 expressing GtfK (lane 3); NM522 with pGSG401 expressing GtfL (lane 4); NM522 with pGSG501 expressing GtfM (lane 5). The number on the left are pH markers. The positions of the bands of activity of GtfM are indicated on the right.

Table 5.4. Electrophoretic Characteristics of *S. salivarius* and Phagemid-Encoded GTFs

Source	pI	Molecular Weight from SDS PAGE	
		Range	Major Band(s)
<i>S. salivarius</i>	4.9, 5.3		137 000 143 000
pGSG101 (GtfJ)	5.1-5.4	134 000 - 163 000	153 000
pGSG201 (GtfK)	5.1-5.4	131 000 - 171 000	140 000 158 000
pGSG401 (GtfL)	5.3	120 000 - 156 000	128 000
pGSG501 (GtfM)	5.4, 9.2	125 000 - 167 000	153 000

results in increased growth of the background *E. coli* creating a halo effect, whereas the formation of a polymer results in plaque opacity. By growing the gene bank of *S. salivarius* chromosomal DNA cloned into λ L47.1 on M9 mineral salts agar with sucrose, clones expressing two GTF activities that differed from the GtfJ and GtfK of *S. salivarius* ATCC 25975 were isolated.

The two GTF activities, GtfL and GtfM, expressed by the bacteriophage- λ clones and the phagemid subclones were primer-independent. In fact, the activity of GtfL was inhibited by the presence of 0.02% (w/v) dextran T-10 (Tables 5.1, 5.2). Neither the *gtfL* nor the *gtfM* gene appeared to be the putative *gtfX* gene which lies upstream of the *gtfJK* locus (see Section 4). Furthermore, the restriction maps of these two genes showed that they did not overlap with each other nor with the *gtfJK* locus. By SDS-PAGE the molecular weight of the largest staining band of GtfM (167 000) was shown to be similar to that of the GtfJ (molecular weight 163 000) and GtfK (molecular weight 171 000). The GtfL on the other hand was slightly smaller, with the largest of the staining bands having a molecular weight 156 000.

Other workers have reported the cloning of primer-independent activities from other oral streptococci (Aoki *et al.*, 1986; Hanada and Kuramitsu, 1988; Russell *et al.*, 1990; Hanada *et al.*, 1991; Fujiwara *et al.*, 1992). The GtfB and GtfC of *S. mutans* GS5 which have been reported to produce both water-soluble and water-insoluble glucan in the absence of primer when expressed in *E. coli*, have been shown to be primer-stimulated (the amount of glucan produced by the enzyme increases in the presence of dextran T-10) when expressed in *S. milleri*. Under these conditions both enzymes produce primarily insoluble glucan (Fukushima *et al.*, 1992). Insertional inactivation has indicated that both the *gtfB* and *gtfC* genes are required for sucrose-dependent colonization by *S. mutans* (Aoki *et al.*, 1986; Hanada and Kuramitsu, 1988) and are important for smooth-surface caries formation (Munro *et al.*, 1991; Yamashita *et al.*, 1993).

Primer-independent GTF-Is such as the GtfL isolated from *S. salivarius* have not been reported in other species of oral streptococci. The GtfL activity differs from the cloned GtfB and GtfC activities of *S. mutans* expressed in *E. coli* in that these enzymes produce both soluble glucan and insoluble glucan with

glucan production stimulated or unaffected by the addition of dextran T-10 (Aoki *et al.*, 1986; Pucci *et al.*, 1987; Hanada and Kuramitsu, 1987). However, an activity similar to that of GtfL has previously been reported from a human oral *S. salivarius* isolate (Sato and Inoue, 1991). The activity was similar to GtfL in that it produced an insoluble glucan, was inhibited by the presence of soluble dextran, and had a similar pI (4.8) and molecular weight (165 000). The glucan product reported by Sato and Inoue (1991) contained high percentages of both α -1,3- and α -1,6-linked glucosyl residues and was susceptible to both commercial α -1,6-dextranase and α -1,3-glucanase. This contrasts with the glucan produced in this study by GtfL which was not digested by *Chaetomium gracile* α -1,6-dextranase.

Primer-independent GTF-Ss have been cloned from two other species of oral streptococci. They are the *gtfS* of *S. downei* and the *gtfT* of *S. sobrinus* (Gilpin *et al.*, 1985; Hanada *et al.*, 1991). *S. sobrinus* is known to produce another primer-independent GTF-S, with an activity similar to that of the GtfS of *S. downei* (Takehara *et al.*, 1984). Although the glucans produced by these two primer-independent GTFs possess primarily α -1,6-linked glucosyl residues, they differ with regard to their chain length, and the nature and length of their side chains. The GTF-S (similar to the GtfS of *S. downei*) of *S. sobrinus* produces short chain oligo-isomaltosaccharides, whereas the product of the GtfT of *S. sobrinus* is a high molecular weight α -1,6-glucan with a high percentage of α -1,3-branch points (Taylor *et al.*, 1990; Hanada *et al.*, 1991; Hanada *et al.*, 1993). The glucan produced by GtfM is also primarily α -1,6-linked as indicated by its digestion with *Chaetomium gracile* α -1,6-dextranase.

The soluble glucans produced by these GTFs are thought to act as primers for the primer-dependent activities. For instance, the high-molecular weight glucan produced by GtfT is a good primer for the primer-dependent GTFs produced by *S. sobrinus* (McCabe, 1985), whereas insertional inactivation of the *gtfS* of *S. downei* prevents the formation of water-insoluble glucan and reduces the ability of the mutant to adhere to glass (Gilmore *et al.*, 1993). However, a similar GTF of *S. sobrinus*, which also produces oligo-isomaltosaccharides was a poor primer for the GTF-I of *S. sobrinus* (Yamashita *et al.*, 1989). Interestingly, the combined action of the primer-independent GTF-S and the primer-dependent

GTF-I of *S. sobrinus* OMZ 176 produces an adherent glucan (Inoue *et al.*, 1982) which allows sucrose-dependent colonization of mutans streptococci. In contrast insoluble glucan produced by GTF-I in the presence of dextran T-10 is non-adherent.

In the following Chapters (6-8) further analysis of the GtfL and GtfM of *S. salivarius* and their respective glucans is described. The GtfL was of particular interest as it appeared to differ from any other GTF described to date, since it produced an insoluble glucan in the absence of primer.

CHAPTER 6

SEQUENCE ANALYSIS OF THE *gtfL* AND *gtfM* GENES

6.1 INTRODUCTION

Several genes encoding for extracellular GTFs have now been cloned and expressed in *E. coli* and the sequences of nine of these *gtf* genes have been published (see section 1.5). Analysis of the deduced amino acid sequences of these genes has revealed that all of them code for proteins sharing a common domain structure. A signal sequence, similar to those found in other streptococcal extra-cytoplasmic proteins (Ferretti *et al.*, 1987) is followed by a highly variable region of approximately 200 amino acids that aligns extremely poorly (Giffard *et al.*, 1993; see below). This region is followed by one that has been designated the 'catalytic domain'. This region encompasses the majority of the N-terminal two-thirds of the enzyme and is responsible for sucrose hydrolysis. An active site aspartic acid is contained within this domain and site-directed mutagenesis of this amino acid has been shown to inhibit GTF activity (Mooser *et al.*, 1991; Kato *et al.*, 1992).

The C-terminal one-third enables the GTFs to bind to glucan molecules (Ferretti *et al.*, 1987; Mooser and Wong, 1988; Kobayashi *et al.*, 1989; Wong *et al.*, 1990). These glucan-binding domains (GBDs) and the closely related glucan-binding protein (GBP) of *S. mutans* are largely composed of a series of repeated sequences (Shiroza *et al.*, 1987; Ferretti *et al.*, 1987; Ueda *et al.*, 1988; Gilmore *et al.*, 1990; Banas *et al.*, 1990; Honda *et al.*, 1990; Abo *et al.*, 1991; Giffard *et al.*, 1991, 1993). As a result, this region also aligns poorly due to the large number of duplications and rearrangements that have occurred in individual proteins (Giffard *et al.*, 1993; Giffard and Jacques, 1994).

Recently, a relationship between these repeats and those of the ligand-binding repeats of the toxins A and B of *C. difficile* as well as the PspA protein and certain lysins from *S. pneumoniae* has been observed both immunologically and at the amino acid sequence level (Gilmore *et al.*, 1990; Wren, 1991; Wren *et al.*, 1991; Yother and Briles, 1992), suggesting that sequences of this type may play

a role in the biology and/or pathogenicity of a range of Gram-positive bacteria. The method of analysing these repeat sequences has in the main been addressed by classifying them into different classes on the basis of high levels of sequence similarity. For example, in the streptococcal GTFs and GBP, 'A', 'B', 'C', and 'D' classes of repeats have been defined (Ferretti *et al.*, 1987; Gilmore *et al.*, 1990; Banas *et al.*, 1990; Abo *et al.*, 1991; Giffard *et al.*, 1991, 1993).

Two recent analyses have extended these previous studies on repeat structures (von Eichel-Streiber *et al.*, 1992; Giffard and Jacques, 1994). The latest of these analyses (Giffard and Jacques, 1994) has defined a fundamentally conserved repeating motif, designated the 'YG' motif, that transcends the boundaries of all previously designated repeat structures found in the related carboxyl-terminal ligand-binding domains of streptococcal GTFs, GBP, *C. difficile* toxin A and *S. pneumoniae* autolysin as well as the carboxyl-terminus of *S. pneumoniae* PspA protein. While it was shown that in any given GTF, pairs of these motifs may possess a wide range of similarity, certain residues are highly conserved. The wide range of similarity indicates that multiple duplication events must have taken place throughout the evolutionary history of the genes, while the conserved residues suggest that a process of functional selection is paramount and is still operative, despite the extensive genetic rearrangements that have taken place since the origin of the ancestral sequence.

Despite the sequences of nine *gtf* genes being known, no obvious explanation for primer-dependence and/or the nature of the linkages in the glucan product synthesized from sucrose by the enzymes coded by these genes have been forthcoming from these analyses. In the case of *S. salivarius* ATCC 25975 two *gtf* genes, *gtfJ* and *gtfK* have been sequenced (Giffard *et al.*, 1991; Giffard *et al.*, 1993). Although the enzymes encoded by these genes produce different glucans (Pitty *et al.*, 1989) both activities were considered to be primer-dependent. As a result, it was hoped that the sequence comparison of the deduced amino acid sequences of the genes encoding primer-independent GTFs with those for primer-dependent GTFs from the same species (analysed in conjunction with or without other similarly primer-dependent or primer-independent GTFs from other species) might allow the identification of regions of the enzymes which conferred primer-

dependency.

6.2 METHODS

6.2.1 DNA Sequence Determination

All sequence determinations were carried out using CsCl purified double stranded templates (2.4.4) via the chain-termination method (Sanger *et al.*, 1977; Section 2.4.18). Sequencing was carried out in both directions using custom-synthesized oligonucleotides. The DNA that was to be sequenced was first subcloned from λ D-10 or λ C-13 into either of the *flori*-containing phagemids pIBI30 or pIBI31 (Table 2.1). The *gtfL* gene was sequenced using the subcloned 4.8kbp and 4.1kbp *EcoRI* fragments (pGSG403 and pGSG404 respectively), with the junction between these fragment being sequenced on the 6.2kbp *Bam*HI/*Xba*I fragment in pGSG401 (Fig. 6.1A). The position of the *gtfM* gene was determined by sequencing from the *EcoRI* site of the 4.8kbp *EcoRI/XbaI* fragment in pGSG502 (Fig. 6.1B). Subsequent sequencing was then carried out using the subcloned 5.3kbp *XbaI* fragment (pGSG501). All sequencing reactions were carried out using the Pharmacia T7 sequencing kit according to the manufacturer's instructions using [³⁵S]-dATP (Dupont NEN). Sequences were assembled using the IBI Pustell sequence analysis software version 2.03. Predicted signal sequence cleavage sites were determined using the 'sigseq' program accessed via the ANGIS facility at the University of Sydney. This program is based on the predictions of von Heijne (1983).

6.2.2 Multiple Alignments and Phylogenetic Tree Construction

Translated open reading frames of *gtf* genes, with the exception of those generated within this laboratory, were downloaded from Genbank using the ANGIS facility at the University of Sydney. All eleven GTFs were aligned using the 'ClustalV' program (Higgins *et al.*, 1992) using default parameters. A region of this alignment which was highly conserved was selected for further analysis. Omission of domains which aligned poorly enabled a multiple alignment with a relatively small number of gaps to be constructed. Phylogenetic analyses of the

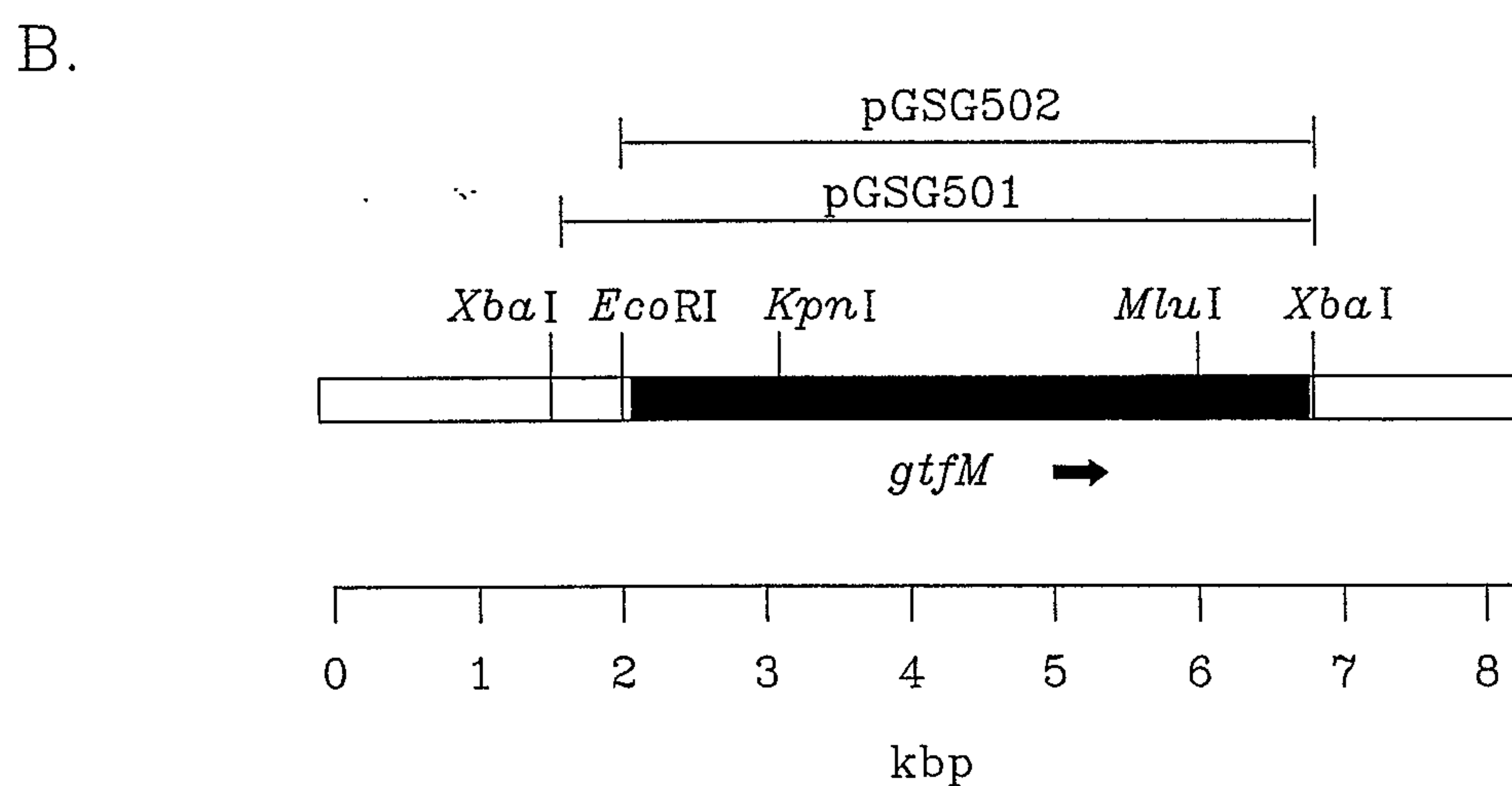
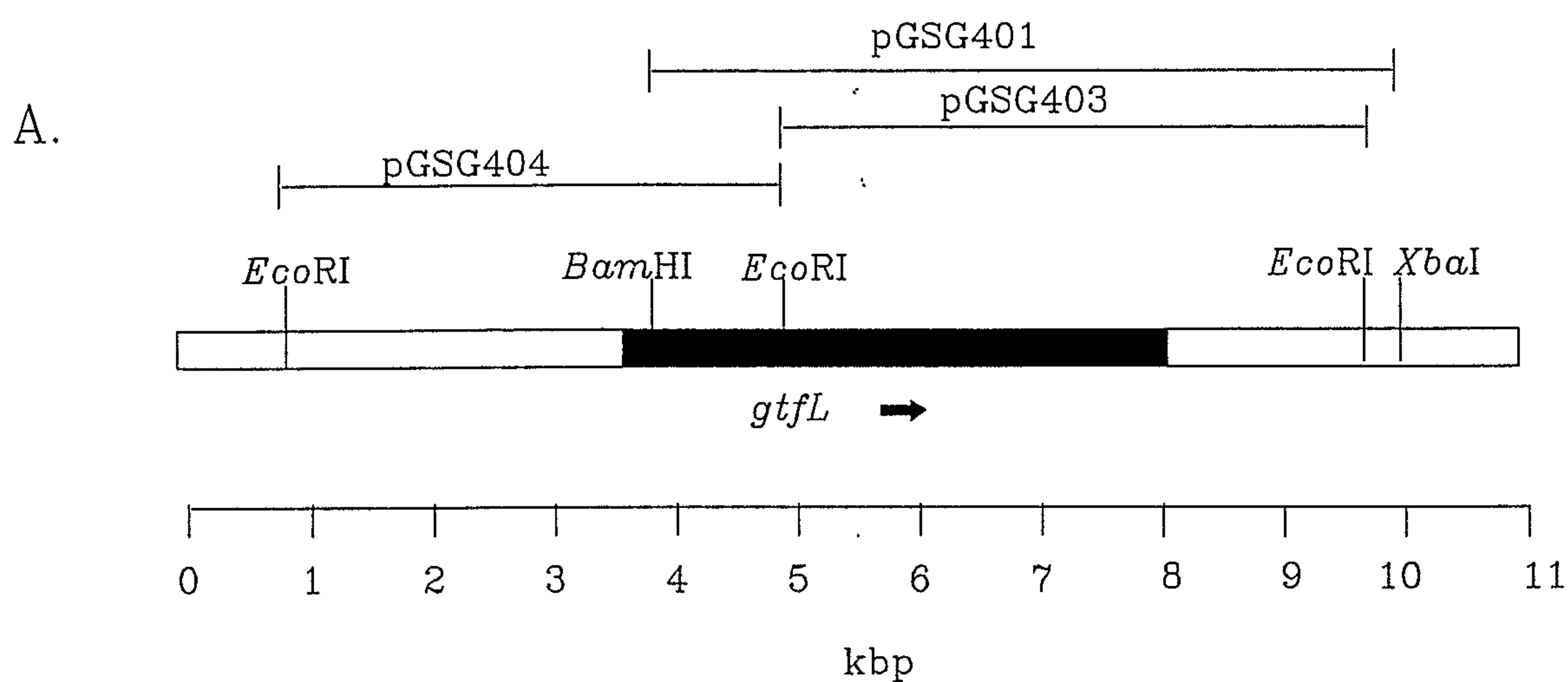


Fig. 6.1. Endonuclease restriction maps of the λ D-10 encoding *GtfL* (A), and λ C-13 encoding *GtfM* (B). The subclones used for sequencing, the position of sequenced *gtf* genes (solid bars) and the direction of transcription (arrows) are also shown.

large conserved region utilized the 'ClustalV' program in conjunction with the 'PHYLP' (Phylogeny Inference Package) Version 3.5c package of software written by Joseph Felsenstein accessed via the ANGIS facility at the University of Sydney. For this analysis gap penalties were modified in order to obtain the best alignment possible (Giffard *et al.*, 1993). The alignment in 'ClustalV' used a pair gap and multiple fixed gap penalty of 10, while the floating gap penalty was set at 30 and a PAM 250 matrix was used. The input to 'PHYLP' was obtained from the 'ClustalV' program and sequentially processed through the programs, 'protdist', 'fitsch' and 'drawtree'.

Phylogenetic analysis of the 'YG' repeats used 'Clustal V' default parameters as this resulted in no gaps. In this instance, the 'Clustal V' output was processed through the 'PHYLP' programs 'protdist', 'fitsch' and 'drawgram'.

6.3 RESULTS AND DISCUSSION

6.3.1 Analysis of *gtfL*

6.3.1.1 Nucleotide Sequence

The nucleotide sequence of the *gtfL* gene was determined in both directions and is shown in Fig. 6.2. The sequence confirmed the restriction map previously described (Section 5). The *gtfL* gene is encoded by a 4470bp open reading frame preceded by a putative ribosome-binding-site sequence (AGGAG) located 10bp upstream of a methionine codon. A potential promoter was identified with a -10 (TATAAAT) at base 103 and a -35 (TTCAATG) at base 82 (Fig. 6.2).

6.3.1.2 Deduced Amino Acid Sequence

The *gtfL* gene has two possible ATG initiation codons (positions 108 and 132). However, as GTFs are secreted enzymes they are expected to possess a signal sequence. In the case of *S. downei*, N-terminal sequence analysis of the mature secreted GtfI (Ferretti *et al.*, 1987) has shown that the proposed signal sequence does in fact exist and is cleaved from the mature protein at the position predicted by von Heijne's '-3, -1 rule' (von Heijne, 1983). Based on the positioning of a

Fig. 6.2. Nucleotide sequence and deduced amino acid sequence of *gtfL*. A potential ribosome binding site (RBS), -10 and -35 promoter elements and signal sequence cleavage site (arrow) are indicated.

10	20	30	40	50	60	
TCAAAACTA ATATTATTAA AACAAACAAA GGACTTTTCT AAAAAATATG AAAATTTATC						
70	80	90	100	110		
GAATCTTATT GAAAAAAGC <u>TTTCAATGAT</u> TTAGAATGAA <u>GCTATAA</u> <u>ATG</u> GGA AAA <u>AGG</u>						
		-35		-10	RBS	
120	130	140	150	160	170	
<u>AGT</u> CTT TTT ATT ATG GAC AAA AAA GTA CAT TAT AAA ATG CAC AAA GTT AAG						
		Met Asp Lys Lys Val His Tyr Lys Met His Lys Val Lys				
		gtfL→				
	180	190	200	210	220	
AAA CAA TGG GTA ACC ATC GCA GTG ACA GGT TTG TCT CTC GGT GCA GTC TCA						
	Lys Gln Trp Val Thr Ile Ala Val Thr Gly Leu Ser Leu Gly Ala Val Ser					
	230	240	250	260	270	
GCG GTT TCT TTA GGA ACT AAC GAC GGT GTC GTT CAG GCA GAC GAA CAT ACG						
	Ala Val Ser Leu Gly Thr Asn Asp Gly Val Val Gln Ala Asp Glu His Thr					
		↑ signal sequence cleavage site				
	280	290	300	310	320	
GAT GCT ACT GTA GCC ATC CCC GAT ATT ACT GTT GAT ACG GGT ACT GTA TCT						
	Asp Ala Thr Val Ala Ile Pro Asp Ile Thr Val Asp Thr Gly Thr Val Ser					
	330	340	350	360	370	
AAC GAT ACA ACG GCT GCT CAG GAT CCA ACT ACA GCA GTT GCA GCG ACA AAC						
	Asn Asp Thr Thr Ala Ala Gln Asp Pro Thr Thr Ala Val Ala Ala Thr Asn					
	380	390	400	410	420	
GAT GTC GCA ACT GAC CAA GCA ACT CCA ACA GCA ACT TTC GAT TTG ACA ACT						
	Asp Val Ala Thr Asp Gln Ala Thr Pro Thr Ala Thr Phe Asp Leu Thr Thr					
	430	440	450	460	470	
GAT ACA ACA AAT ACA GTT GCA GCG AAT GCT GTT GAT ACG GTA GCG ACT GTC						
	Asp Thr Thr Asn Thr Val Ala Ala Asn Ala Val Asp Thr Val Ala Thr Val					
	480	490	500	510	520	
GGT ACT GAT CGT GCG GCA ACT ACA AAC GAT ACA ACA GCA ACC AAT GAT ACA						
	Gly Thr Asp Arg Ala Ala Thr Thr Asn Asp Thr Thr Ala Thr Asn Asp Thr					
	530	540	550	560	570	
GCT GTT GAT ACA ACA AAT AAC AAT ACT ACA ACT GAC ACA ACA ACA GTT ACT						
	Ala Val Asp Thr Thr Asn Asn Asn Thr Thr Thr Asp Thr Thr Thr Val Thr					
	580	590	600	610	620	
GAT CGC GCA GCG ACA ACT GAA CGT CGT GCG ACT GGT GCT CGC CGT GGT CCT						
	Asp Arg Ala Ala Thr Thr Glu Arg Arg Ala Thr Gly Ala Arg Arg Gly Pro					
	630	640	650	660	670	680
ACA GGA GGC CGT CGT GCT ACC CCA GTT AAC GGC AAT ACA AAT AAT GCC AAC						
	Thr Gly Gly Arg Arg Ala Thr Pro Val Asn Gly Asn Thr Asn Asn Ala Asn					
	690	700	710	720	730	
AAT ACT GTC ACA GTA GTA AAT AAC GAC CTT CCA GCA ACT AAT AAT GTC GTA						
	Asn Thr Val Thr Val Val Asn Asn Asp Leu Pro Ala Thr Asn Asn Val Val					
	740	750	760	770	780	
ACT GAT GGT CCA AGC CAC ATC AAA ACT ATT AAT GGA AAA CAA TAC TAC GTT						
	Thr Asp Gly Pro Ser His Ile Lys Thr Ile Asn Gly Lys Gln Tyr Tyr Val					
	790	800	810	820	830	
GAA GAT GAT GGC ACT ATT CGT AAG AAC TAT GTT CTT GAA CGT ATC GGT GGA						
	Glu Asp Asp Gly Thr Ile Arg Lys Asn Tyr Val Leu Glu Arg Ile Gly Gly					

840	850	860	870	880
AGC CAA TAC TTT AAC GCT GAA ACT GGT GAA TTG TCT AAC CAA AAA GAA TAC				
Ser Gln Tyr Phe Asn Ala Glu Thr Gly Glu Leu Ser Asn Gln Lys Glu Tyr				
890	900	910	920	930
CGT TTC GAT AAG AAC GGT GGA ACA GGT TCA TCA GCT GAC AGC ACA AAC ACA				
Arg Phe Asp Lys Asn Gly Gly Thr Gly Ser Ser Ala Asp Ser Thr Asn Thr				
940	950	960	970	980
AAC GTT ACT GTA AAT GGT GAT AAA AAT GCC TTT TAC GGT ACA ACT GAT AAA				
Asn Val Thr Val Asn Gly Asp Lys Asn Ala Phe Tyr Gly Thr Thr Asp Lys				
990	1000	1010	1020	1030
GAT ATC GAA CTT GTT GAT GGT TAC TTC ACT GCC AAC ACA TGG TAC CGT CCA				
Asp Ile Glu Leu Val Asp Gly Tyr Phe Thr Ala Asn Thr Trp Tyr Arg Pro				
1040	1050	1060	1070	1080
AAA GAA ATC TTG AAA GAT GGT AAA GAG TGG ACT GCC TCA ACT GAA AAT GAC				
Lys Glu Ile Leu Lys Asp Gly Lys Glu Trp Thr Ala Ser Thr Glu Asn Asp				
1090	1100	1110	1120	1130
AAA CGT CCA TTG CTT ACA GTT TGG TGG CCA AGC AAG GCT ATC CAA GCG AGC				
Lys Arg Pro Leu Leu Thr Val Trp Trp Pro Ser Lys Ala Ile Gln Ala Ser				
1140	1150	1160	1170	1180
TAC CTC AAC TAC ATG AAG GAG CAA GGT CTC GGT ACC AAT CAG ACC TAT ACG				
Tyr Leu Asn Tyr Met Lys Glu Gln Gly Leu Gly Thr Asn Gln Thr Tyr Thr				
1200	1210	1220	1230	1240
TCA TTC AGC TCA CAA ACC CAA ATG GAC CAA GCA GCA CTT GAA GTT CAA AAG				
Ser Phe Ser Ser Gln Thr Gln Met Asp Gln Ala Ala Leu Glu Val Gln Lys				
1250	1260	1270	1280	1290
CGT ATT GAA GAA CGT ATT GCG CGT GAA GGT AAC ACT GAC TGG TTG CGT ACA				
Arg Ile Glu Glu Arg Ile Ala Arg Glu Gly Asn Thr Asp Trp Leu Arg Thr				
1300	1310	1320	1330	1340
ACT ATC AAA AAT TTC GTG AAG ACC CAA CCA GGT TGG AAT TCA ACT TCT GAA				
Thr Ile Lys Asn Phe Val Lys Thr Gln Pro Gly Trp Asn Ser Thr Ser Glu				
1350	1360	1370	1380	1390
AAC CTT GAT AAT AAT GAC CAC TTG CAA GGT GGT GCC CTT CTC TAT AAC AAC				
Asn Leu Asp Asn Asn Asp His Leu Gln Gly Gly Ala Leu Leu Tyr Asn Asn				
1400	1410	1420	1430	1440
GAT AGT CGT ACA AGC CAT GCT AAC TCA GAT TAT CGT CTT TTG AAC CGC ACA				
Asp Ser Arg Thr Ser His Ala Asn Ser Asp Tyr Arg Leu Leu Asn Arg Thr				
1450	1460	1470	1480	1490
CCA ACT AGT CAA ACA GGT AAA CAC AAT CCT AAG TAT ACT AAG GAT ACA AGT				
Pro Thr Ser Gln Thr Gly Lys His Asn Pro Lys Tyr Thr Lys Asp Thr Ser				
1500	1510	1520	1530	1540
AAC GGT GGT TTC GAG TTC TTG TTG GCA AAT GAT ATT GAC AAC TCA AAT CCA				
Asn Gly Gly Phe Glu Phe Leu Leu Ala Asn Asp Ile Asp Asn Ser Asn Pro				
1550	1560	1570	1580	1590
GCG GTT CAG GCT GAG CAG TTG AAC TGG CTC CAC TAC ATC ATG AAC ATC GGT				
Ala Val Gln Ala Glu Gln Leu Asn Trp Leu His Tyr Ile Met Asn Ile Gly				
1600	1610	1620	1630	1640
ACT ATT ACT GGT GGT TCT GAA GAT GAA AAC TTC GAT GGT GTT CGT GTC GAT				
Thr Ile Thr Gly Gly Ser Glu Asp Glu Asn Phe Asp Gly Val Arg Val Asp				

1650	1660	1670	1680	1690	1700
GCC GTG GAT AAT GTC AAT GCT GAC CTT CTC CAA ATT GCT TCA GAT TAT TTC					
Ala Val Asp Asn Val Asn Ala Asp Leu Leu Gln Ile Ala Ser Asp Tyr Phe					
	1710	1720	1730	1740	1750
AAA GCT AAA TAT GGT GCT GAT CAA AGT CAA GAT CAA GCC ATT AAA CAC TTG					
Lys Ala Lys Tyr Gly Ala Asp Gln Ser Gln Asp Gln Ala Ile Lys His Leu					
	1760	1770	1780	1790	1800
TCT ATC TTG GAA GCA TGG TCT CAC AAC GAT GCC TAT TAC AAT GAA GAC ACT					
Ser Ile Leu Glu Ala Trp Ser His Asn Asp Ala Tyr Tyr Asn Glu Asp Thr					
	1810	1820	1830	1840	1850
AAA GGT GCC CAA TTG CCA ATG GAT GAC CCT ATG CAC TTA GCA TTG GTA TAT					
Lys Gly Ala Gln Leu Pro Met Asp Asp Pro Met His Leu Ala Leu Val Tyr					
	1860	1870	1880	1890	1900
TCA CTC TTG CGT CCA ATT GGT AAC CGC TCA GGG GTT GAA CCA CTC ATT TCA					
Ser Leu Leu Arg Pro Ile Gly Asn Arg Ser Gly Val Glu Pro Leu Ile Ser					
	1910	1920	1930	1940	1950
AAT AGT TTG AAT GAC CGT TCT GAA AGT GGT AAA AAC TCT AAA CGC ATG GCT					
Asn Ser Leu Asn Asp Arg Ser Glu Ser Gly Lys Asn Ser Lys Arg Met Ala					
	1960	1970	1980	1990	2000
AAC TAC GCC TTT GTT CGT GCT CAC GAT TCA GAG GTA CAA TCA ATC ATT GGT					
Asn Tyr Ala Phe Val Arg Ala His Asp Ser Glu Val Gln Ser Ile Ile Gly					
	2010	2020	2030	2040	2050
CAA ATC ATC AAA AAT GAA ATC AAC CCT CAA TCA ACA GGG AAT ACC TTC ACG					
Gln Ile Ile Lys Asn Glu Ile Asn Pro Gln Ser Thr Gly Asn Thr Phe Thr					
2060	2070	2080	2090	2100	
CTT GAT GAA ATG AAG AAA GCC TTT GAA ATC TAC AAC AAA GAC ATG CGT AGT					
Leu Asp Glu Met Lys Lys Ala Phe Glu Ile Tyr Asn Lys Asp Met Arg Ser					
2110	2120	2130	2140	2150	
GCT AAC AAG CAA TAC ACA CAA TAT AAC ATC CCA TCA GCA TAT GCC TTG ATG					
Ala Asn Lys Gln Tyr Thr Gln Tyr Asn Ile Pro Ser Ala Tyr Ala Leu Met					
2160	2170	2180	2190	2200	2210
CTT ACG CAT AAG GAC ACT GTG CCA CGT GTT TAC TAT GGA GAC ATG TAC ACA					
Leu Thr His Lys Asp Thr Val Pro Arg Val Tyr Tyr Gly Asp Met Tyr Thr					
	2220	2230	2240	2250	2260
GAT GAT GGT CAA TAT ATG GCT CAA AAA TCA CCA TAT TAT GAT GCT ATC GAA					
Asp Asp Gly Gln Tyr Met Ala Gln Lys Ser Pro Tyr Tyr Asp Ala Ile Glu					
	2270	2280	2290	2300	2310
ACC CTC TTG AAA GGC CGT ATC CGC TAT GCA GCC GGT GGT CAA GAT ATG AAG					
Thr Leu Leu Lys Gly Arg Ile Arg Tyr Ala Ala Gly Gly Gln Asp Met Lys					
	2320	2330	2340	2350	2360
GTT AAC TAC ATC GGT TAT GGT AAT ACT AAT GGT TGG GAT GCT GCA GGT GTC					
Val Asn Tyr Ile Gly Tyr Gly Asn Thr Asn Gly Trp Asp Ala Ala Gly Val					
	2370	2380	2390	2400	2410
TTG ACA TCT GTA CGT TAT GGT ACA GGG GCT AAC AGT GCT TCT GAT ACT GGT					
Leu Thr Ser Val Arg Tyr Gly Thr Gly Ala Asn Ser Ala Ser Asp Thr Gly					
	2420	2430	2440	2450	2460
ACG GCA GAG ACA CGT AAC CAA GGT ATG GCA GTT ATC GTT TCA AAC CAA CCT					
Thr Ala Glu Thr Arg Asn Gln Gly Met Ala Val Ile Val Ser Asn Gln Pro					

2470	2480	2490	2500	2510
GCC CTT CGT TTG ACA AGC AAC TTG ACA ATT AAC ATG GGT GCG GCT CAC CGT				
Ala Leu Arg Leu Thr Ser Asn Leu Thr Ile Asn Met Gly Ala Ala His Arg				
2520	2530	2540	2550	2560
AAC CAA GCT TAT CGT CCA TTG CTT TTG ACA ACA AAC GAT GGT GTG GCT ACT				
Asn Gln Ala Tyr Arg Pro Leu Leu Leu Thr Thr Asn Asp Gly Val Ala Thr				
2570	2580	2590	2600	2610
TAC TTG AAT GAC AGC GAT GCT AAC GGT ATT GTT AAA TAC ACT GAT GGC AAC				
Tyr Leu Asn Asp Ser Asp Ala Asn Gly Ile Val Lys Tyr Thr Asp Gly Asn				
2620	2630	2640	2650	2660
GGT AAT TTG ACA TTT AGT GCA AAT GAA ATT CGT GGT ATC CGT AAC CCT CAA				
Gly Asn Leu Thr Phe Ser Ala Asn Glu Ile Arg Gly Ile Arg Asn Pro Gln				
2670	2680	2690	2700	2710
GTT GAT GGT TAC CTT GCC GTT TGG GTA CCA GTA GGT GCT TCT GAA AAT CAA				
Val Asp Gly Tyr Leu Ala Val Trp Val Pro Val Gly Ala Ser Glu Asn Gln				
2730	2740	2750	2760	2770
GAC GTG CGT GTA GCA CCA AGT AAA GAA AAG AAC AGC AGT GGT TTG GTT TAC				
Asp Val Arg Val Ala Pro Ser Lys Glu Lys Asn Ser Ser Gly Leu Val Tyr				
2780	2790	2800	2810	2820
GAA TCA AAT GCT GCC CTT GAT TCA CAA GTT ATC TAC GAA GGT TTC TCT AAC				
Glu Ser Asn Ala Ala Leu Asp Ser Gln Val Ile Tyr Glu Gly Phe Ser Asn				
2830	2840	2850	2860	2870
TTC CAA GAC TTC GTT CAA AAT CCA TCA CAA TAC ACT AAT AAG AAG ATT GCT				
Phe Gln Asp Phe Val Gln Asn Pro Ser Gln Tyr Thr Asn Lys Lys Ile Ala				
2880	2890	2900	2910	2920
GAA AAT GCA AAC CTC TTC AAA TCA TGG GGT ATC ACA TCA TTC GAA TTT GCA				
Glu Asn Ala Asn Leu Phe Lys Ser Trp Gly Ile Thr Ser Phe Glu Phe Ala				
2930	2940	2950	2960	2970
CCA CAG TAC GTT TCA AGT GAT GAC GGT AGC TTC CTT GAC TCA GTA ATC CAA				
Pro Gln Tyr Val Ser Ser Asp Asp Gly Ser Phe Leu Asp Ser Val Ile Gln				
2980	2990	3000	3010	3020
AAT GGT TAT GCC TTT ACA GAC CGT TAT GAT ATT GGT ATG AGT AAG GAC AAT				
Asn Gly Tyr Ala Phe Thr Asp Arg Tyr Asp Ile Gly Met Ser Lys Asp Asn				
3030	3040	3050	3060	3070
AAG TAC GGT TCA TTG GCT GAC TTG AAA GCA GCT CTT AAA TCA CTC CAC GCT				
Lys Tyr Gly Ser Leu Ala Asp Leu Lys Ala Ala Leu Lys Ser Leu His Ala				
3080	3090	3100	3110	3120
GTA GGT ATC TCA GCG ATT GCT GAC TGG GTT CCA GAC CAA ATC TAC AAC CTT				
Val Gly Ile Ser Ala Ile Ala Asp Trp Val Pro Asp Gln Ile Tyr Asn Leu				
3130	3140	3150	3160	3170
CCA GGT GAC GAA GTG GTT ACT GCA ACT CGT GTT AAT AAC TAT GGT GAA ACT				
Pro Gly Asp Glu Val Val Thr Ala Thr Arg Val Asn Asn Tyr Gly Glu Thr				
3180	3190	3200	3210	3220
AAA GAT GGT GCA ATC ATC GAT CAT AGC CTC TAC GCT GCT AAA ACA CGT ACC				
Lys Asp Gly Ala Ile Ile Asp His Ser Leu Tyr Ala Ala Lys Thr Arg Thr				
3240	3250	3260	3270	3280
TTT GGT AAT GAC TAC CAA GGT AAA TAT GGT GGT GCC TTC CTT GAT GAA TTG				
Phe Gly Asn Asp Tyr Gln Gly Lys Tyr Gly Gly Ala Phe Leu Asp Glu Leu				

3290	3300	3310	3320	3330
AAA CGT CTC TAC CCA CAA ATC TTT GAC CGT GTT CAA ATC TCT ACT GGT AAA				
Lys Arg Leu Tyr Pro Gln Ile Phe Asp Arg Val Gln Ile Ser Thr Gly Lys				
3340	3350	3360	3370	3380
CGT ATG ACT ACT GAC GAG AAG ATT ACA CAG TGG TCT GCT AAA TAC ATG AAT				
Arg Met Thr Thr Asp Glu Lys Ile Thr Gln Trp Ser Ala Lys Tyr Met Asn				
3390	3400	3410	3420	3430
GGT ACT AAC ATT CTT GAC CGT GGC TCA GAG TAC GTA CTT AAG AAT GGC CTC				
Gly Thr Asn Ile Leu Asp Arg Gly Ser Glu Tyr Val Leu Lys Asn Gly Leu				
3440	3450	3460	3470	3480
AAC GGT TAC TAC GGT ACA AAT GGC GGT AAG GTT AGC CTT CCA AAA GTA GTA				
Asn Gly Tyr Tyr Gly Thr Asn Gly Gly Lys Val Ser Leu Pro Lys Val Val				
3490	3500	3510	3520	3530
GGA AGC AAC CAA TCA ACA AAT GGA GAT AAT CAA AAT GGT GAC GGC AGT GGT				
Gly Ser Asn Gln Ser Thr Asn Gly Asp Asn Gln Asn Gly Asp Gly Ser Gly				
3540	3550	3560	3570	3580
AAG TTT GAA AAA CGT CTC TTC TCT GTA AGA TAC CGT TAT AAC AAT GGC CAA				
Lys Phe Glu Lys Arg Leu Phe Ser Val Arg Tyr Arg Tyr Asn Asn Gly Gln				
3590	3600	3610	3620	3630
TAT GCT AAG AAT GCC TTC ATC AAA GAT AAC GAC GGA AAT GTT TAC TAC TTT				
Tyr Ala Lys Asn Ala Phe Ile Lys Asp Asn Asp Gly Asn Val Tyr Tyr Phe				
3640	3650	3660	3670	3680
GAT AAC TCA GGT CGC ATG GCA GTC GGA GAA AAG ACT ATC GAT GGT AAA CAA				
Asp Asn Ser Gly Arg Met Ala Val Gly Glu Lys Thr Ile Asp Gly Lys Gln				
3690	3700	3710	3720	3730
TAC TTC TTC CTC GCT AAT GGT GTG CAA TTG CGT GAT GGT TAT CGT CAA AAC				
Tyr Phe Phe Leu Ala Asn Gly Val Gln Leu Arg Asp Gly Tyr Arg Gln Asn				
3750	3760	3770	3780	3790
CGT AGA GGT CAA GTA TTC TAC TAC GAC CAA AAC GGT GTC TTG AAT GCA AAT				
Arg Arg Gly Gln Val Phe Tyr Tyr Asp Gln Asn Gly Val Leu Asn Ala Asn				
3800	3810	3820	3830	3840
GGT AAA CAA GAT CCA AAA CCA GAT AAC AAT AAT AAT GCA TCA GGC CGT AAC				
Gly Lys Gln Asp Pro Lys Pro Asp Asn Asn Asn Asn Ala Ser Gly Arg Asn				
3850	3860	3870	3880	3890
CAA TTC GTA CAA ATT GGT AAT AAC GTT TGG GCA TAC TAT GAT GGT AAT GGT				
Gln Phe Val Gln Ile Gly Asn Asn Val Trp Ala Tyr Tyr Asp Gly Asn Gly				
3900	3910	3920	3930	3940
AAA CGT GTG ACA GGT CAC CAA AAC ATC AAT GGT CAA GAA CTC TTC TTC GAC				
Lys Arg Val Thr Gly His Gln Asn Ile Asn Gly Gln Glu Leu Phe Phe Asp				
3950	3960	3970	3980	3990
AAC AAT GGT GTC CAA GTT AAA GGG CGC ACT GTT AAC GAG AAT GGT GCT ATC				
Asn Asn Gly Val Gln Val Lys Gly Arg Thr Val Asn Glu Asn Gly Ala Ile				
4000	4010	4020	4030	4040
CGT TAC TAT GAT GCT AAC TCA GGT GAG ATG GCT CGT AAC CGA TTC GCA GAA				
Arg Tyr Tyr Asp Ala Asn Ser Gly Glu Met Ala Arg Asn Arg Phe Ala Glu				
4050	4060	4070	4080	4090
ATC GAA CCA GGT GTT TGG GCT TAC TTC AAT AAT GAT GGT ACG GCT GTA AAA				
Ile Glu Pro Gly Val Trp Ala Tyr Phe Asn Asn Asp Gly Thr Ala Val Lys				

4100	4110	4120	4130	4140
GGT TCA CAA AAC ATC AAT GGT CAA GAC CTC TAC TTT GAT CAA AAT GGC CGT				
Gly Ser Gln Asn Ile Asn Gly Gln Asp Leu Tyr Phe Asp Gln Asn Gly Arg				
4150	4160	4170	4180	4190
CAG GTC AAA GGT GCT TTG GCA AAC GTA GAT GGC AAC CTT CGT TAC TAT GAT				
Gln Val Lys Gly Ala Leu Ala Asn Val Asp Gly Asn Leu Arg Tyr Tyr Asp				
4200	4210	4220	4230	4240
GTA AAC TCA GGT GAG CTT TAC CGT AAT CGA TTC CAC GAA ATT GAT GGT AGC				
Val Asn Ser Gly Glu Leu Tyr Arg Asn Arg Phe His Glu Ile Asp Gly Ser				
	4260	4270	4280	4290
	TGG TAT TAC TTC GAT GGT AAT GGT AAT GCT GTC AAA GGC ATG GTT AAT ATC			
	Trp Tyr Tyr Phe Asp Gly Asn Gly Asn Ala Val Lys Gly Met Val Asn Ile			
	4310	4320	4330	4340
	AAT GGT CAA AAC CTC CTC TTT GAT AAC AAT GGT AAA CAA ATC AAA GGT CAT			
	Asn Gly Gln Asn Leu Leu Phe Asp Asn Asn Gly Lys Gln Ile Lys Gly His			
	4360	4370	4380	4390
	TTG GTT CGC GTG AAC GGT GTT GTT CGT TAC TTT GAT CCA AAT TCA GGT GAA			
	Leu Val Arg Val Asn Gly Val Val Arg Tyr Phe Asp Pro Asn Ser Gly Glu			
	4410	4420	4430	4440
	ATG GCT GTC AAC CGT TGG GTT GAA GTG AGT CCA GGA TGG TGG GTT TAC TTT			
	Met Ala Val Asn Arg Trp Val Glu Val Ser Pro Gly Trp Trp Val Tyr Phe			
	4460	4470	4480	4490
	GAC GGT GAA GGC CGA GGT CAA ATT TAA TCGTTTCTA AACTGTATGA ACCGAGACAG			
	Asp Gly Glu Gly Arg Gly Gln Ile End			
	4520	4530	4540	4550
	TAGTCTCGGT TTTTATGGCA TGTTTGGATT TCGTAAGGAT TTGGTGATGA CTGAAACTAG			
	4580	4590	4600	4610
	TGATTTTAAA ACTATTTTGA GTTAGTCATA ATAGTCTGAC AGTGGTCGTT ACATCT			

putative ribosome binding site and the sequence similarity to other GTFs, it was predicted that the GtfL protein sequence begins at the second ATG codon (position 132 in Fig. 6.2). By choosing this second ATG codon a putative signal sequence can be readily defined. This putative signal sequence also readily aligns and is highly conserved with respect to the predicted leader sequences in all other known GTF sequences (see 6.3.3). Thus the deduced amino acid sequence of GtfL defines a protein with a molecular weight of 159 973 and a pI of 5.27. The 'sigseq' program predicted a cleavage site of GtfL after amino acid residue 35 (Fig. 6.2), leaving a mature protein of molecular weight 156 844 with a pI of 5.19. The pI of GtfL is similar to that of GtfJ and GtfK of *S. salivarius* ATCC 25975 (Pitty *et al.*, 1989) and is therefore not likely to be distinguished from these activities in *S. salivarius* culture fluid on isoelectric focusing gels. This is consistent with the finding on the isoelectric focusing gel in Chapter 5 (Fig. 5.8), where the pI values of the Gtfs expressed by the λ -clones λ A-8, λ A-33 and λ D-10 appeared identical. The molecular weight of the predicted protein is similar to that of other sequenced GTFs, which range from 147 408 (GtfS of *S. downei*) to 177 706 (GtfI of *S. downei*), and is smaller than the GtfJ and GtfK of *S. salivarius* as was evident in SDS-PAGE analysis (Fig. 5.7). The subclone used for SDS-PAGE and isoelectric focusing in Chapter 5 contained a 65 amino acid N-terminal deletion. This enzyme would therefore have a predicted molecular weight of 153 142 with a pI of 5.20. These values are similar to the pI of 5.3 and maximum molecular weight of 156 000 experimentally determined for the enzyme expressed by the subclone (Table 5.4).

6.3.1.3 Repeat Units of GtfL

All sequenced *gtf* genes thus far have a region containing direct repeat units in the carboxyl-terminal one third of the molecule, thought to be involved in glucan binding (6.1 above). These repeats, which have previously been divided into four different classes (A,B,C,D) on the basis on sequence similarity, (6.1 above) have recently been shown to be composed of a common structural element, designated the 'YG' repeat (Giffard and Jacques, 1994). Analysis of the C-terminus of GtfL indicated that it possessed 15 'YG' repeats and one partial repeat (Fig. 6.3).

L1	1032	RTFGNDYQGKYGGAF ~~L~~ DELKR	1062
L2	1008	KNGLNGYYGTNGGKVSLPKVV	1118
L3	1140	RLFSVRYRYNNGQYAKNAFIK	1160
L4	1162	NDGNVYYFDNSGRMAVGEKTI	1182
L5	1182	IDGKQYFFLANGVQLRDGYRQ	1202
L6	1204	RRGQVFYYDQNGVLNANGKQD	1224
L7	1243	GNNVWAYYDGNGKRVTGHQNI	1263
L8	1263	INGQELFFDNNGVQVKGRTVN	1283
L9	1285	NGAIRYYDANS ~~G~~ EMARNRFAE	1305
L10	1307	EPGVWAYFNNDGTAVKGSQNI	1327
L11	1327	INGQDLYFDQNGRQVKGALAN	1347
L12	1349	DGNLRYYDVNSGELYRNR ~~F~~ HE	1369
L13	1370	IDGSWYYFDGNGNAVKG ~~M~~ VNI	1390
L14	1390	INGQNLLFDNNGKQIKGHLVR	1410
L15	1412	NGVVR ~~Y~~ FDPN ~~S~~ GEMAVNRWVE	1432
L16	1434	SPGWWVYFDGEGRGQI	1449 (end of protein)

Fig. 6.3. ‘YG’ repeats from the C-terminal domain of GtfL. The numbers to the left and right of the repeats indicate the first and last amino acids of each repeat sequence.