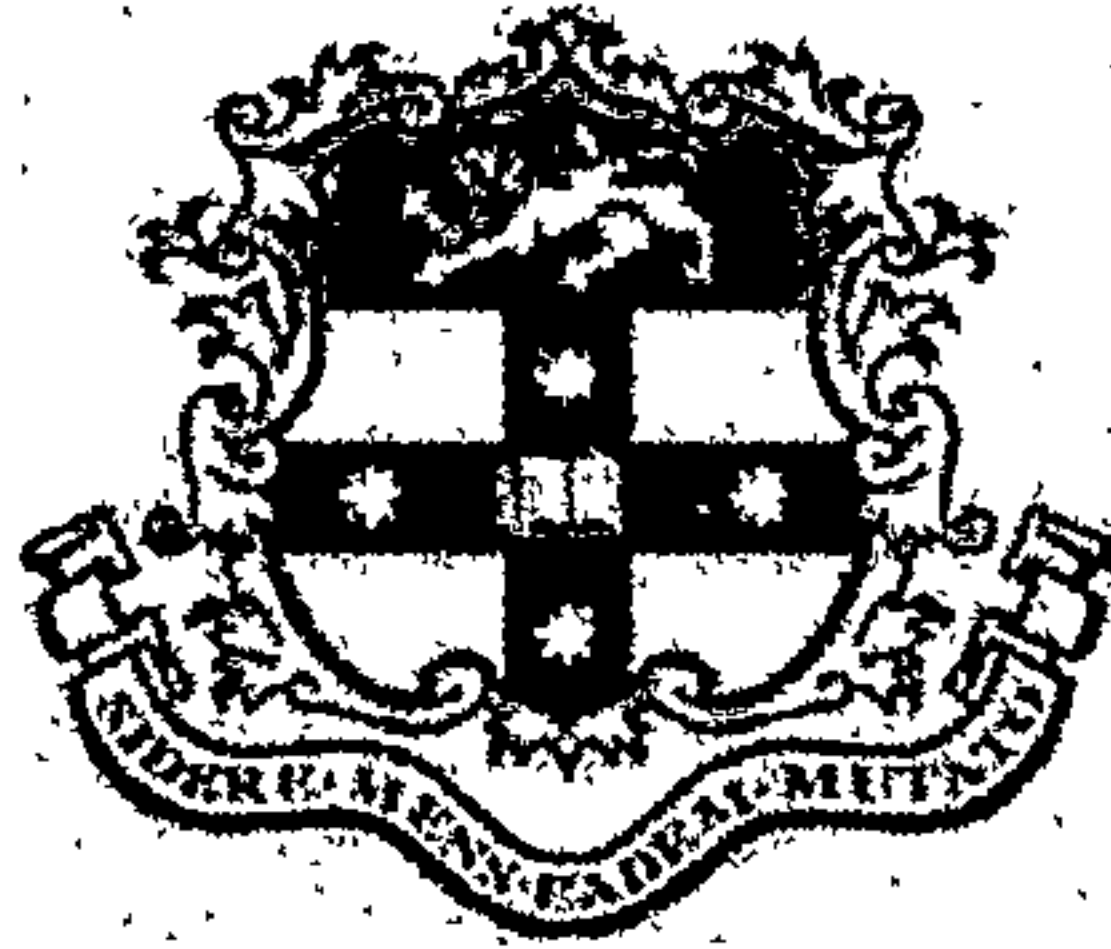

Doctorate of Philosophy
The University of Sydney



The Cell-Associated Fructosyltransferase of
Streptococcus salivarius ATCC 25975

Catherine Rathsam
2001

STATEMENT OF AUTHORSHIP

The research presented in this thesis was undertaken at the Institute of Dental Research (Department of Oral Biology) at the United Dental Hospital of Sydney. The thesis contains no research material which has been submitted for the award of any other degree and, to the best of my knowledge and belief, contains no experimental material previously published or written by another person, except where due reference is given.

Catherine Rathsam

23rd March, 2001

Sydney, Australia

ACKNOWLEDGEMENT

To my supervisor, Dr. Nick Jacques, I would like to express my deepest gratitude for his diligent guidance, and his unmatched compassion and humour.

I would also like to thank Professor Ken Knox and the acting director of the Institute of Dental Research, Dr. Neil Hunter for his kind support and use of the facilities at the Institute. This is extended to friends and colleagues at the Institute who have provided help and friendship; Dr. Derek Harty, Dr. Phil Giffard, Dr. Christine Simpson, Dr. Donna Song, Donna O'Connor, Dr. Mayuri Paramaesvaran and Dr. Ky-Anh Nguyen.

I would also like to express my thanks to my family; my parents, Leander, Van, Jayden and James for their constant support.

I would like to express my appreciation to the Australian National Health and Medical Research Council (NHMRC) for a Biomedical Scholarship.

This work was funded by an NHMRC grant awarded to Dr. Nick Jacques at the Institute of Dental Research.

PREFACE

Some of the work described in this thesis has been reported in the following publications.

Research Publications

Giffard, PM, **CA Rathsam**, E Kwan, DWL Kwan, KM Bunny, S-P Koo and NA Jacques. 1993. The *ftf* gene encoding the cell-bound fructosyltransferase of *Streptococcus salivarius* ATCC 25975 is preceded by an insertion sequence and followed by *FUR1* and *clpP* homologues. *J Gen Microbiol* **139**: 913-920.

Rathsam, C, PM Giffard and NA Jacques. 1993. The cell-bound fructosyltransferase of *Streptococcus salivarius*: The carboxyl-terminus specifies attachment in a *Streptococcus gordonii* model system. *J Bacteriol* **175**: 4520-4527.

Rathsam, C, and NA Jacques. 1997. Development of a technique for multiple site-directed mutagenesis of the *ftf* gene of *Streptococcus salivarius* containing palindromic sequences. *FEMS Microbiol Letts* **153**: 447-453.

Rathsam, C, and NA Jacques. 1998. Role of C-terminal domains in surface attachment of the fructosyltransferase of *Streptococcus salivarius* ATCC 25975. *J Bacteriol* **180**: 6400-6403.

Conference Presentations

Giffard, PM, CP Milward, S-P Koo, **C Rathsam** and NA Jacques. 1990. Does a single gene code for fructosyltransferase activities in *Streptococcus salivarius*? *Aust Microbiologist* **11**: 235.

Rathsam, C, ENO Kwan, PM Giffard and NA Jacques. 1992. Molecular characterization of the *ftf* encoding region of *Streptococcus salivarius*. *J Dent Res* **72**: 733.

Rathsam, C, ENO Kwan, PM Giffard and NA Jacques. 1992. Characterization of the *ftf* of *Streptococcus salivarius*. *Aust Microbiologist* **13**: A69.

Rathsam, C, and NA Jacques. 1994. Fructan synthesis by the oral bacterium, *Streptococcus salivarius*. *Abstr 34th Meeting Aust Soc Plant Physiol*. p163.

Rathsam, C, and NA Jacques. 1996. Development of a technique for site-directed mutagenesis of the *ftf* gene. *Faculty of Dent Res Day*. Westmead Hospital, Sydney.

Rathsam, C, and NA Jacques. 1996. Development of a technique for site-directed mutagenesis of the *ftf* gene. *Aust Microbiologist* **17**: A59.

Rathsam, C, and NA Jacques. 1997. Analysis of the binding of fructosyltransferase to the surface of *Streptococcus salivarius*. *Abstr 4th Aust Conf Molec Analysis Bact Path*. p66.

Song, DD, **C Rathsam** and NA Jacques. 1999. Molecular and kinetic analysis of the β -D-fructosyltransferase of *Streptococcus salivarius*. *Proc 5th Aust Conf Molec Analysis Bact Path*. p51.

Song, DD, **C Rathsam** and NA Jacques. 1999. Molecular and kinetic analysis of the β -D-fructosyltransferase of *Streptococcus salivarius*. *Abstr XIV Lancefield Int Symp Streptococci Streptococcal Diseases*. P10.19.

SUMMARY

The fructosyltransferases (Ftfs) of oral streptococci polymerize the fructosyl moiety of sucrose to form fructan, and are considered to contribute to the virulence of cariogenic bacteria. Unlike the Ftfs of other bacteria which are secreted, the Ftf of *Streptococcus salivarius* ATCC25975 is initially cell-bound and released in the presence of its substrate sucrose. A Ftf encoding insert isolated from a genomic library of *S. salivarius* was cloned and sequenced. The three open reading frames surrounding the *ftf* gene were found to show strong homology to a uracil phosphoribosyltransferase (UPRTase), a ClpP ATP-dependent protease and an insertion sequence-like element.

The *ftf* gene was analyzed and the amino acid sequence determined. Amino acid sequencing of the native Ftf isolated from *S. salivarius* confirmed the identity of the clone and located two processing sites at the N-terminus of the protein. The deduced amino acid sequence of the Ftf was compared with the secreted Ftfs of other organisms. Several highly conserved residues within the catalytic domain were found, as well as a unique 178 residue proline-glycine- and threonine-serine-rich extension at the C-terminus of the protein. This region was followed by a 25 amino acid hydrophobic domain and a small charged tail. Structurally, this arrangement shows strong similarity to C-terminal regions of surface-bound proteins of other Gram-positive bacteria. However, the LPXTG wall anchoring motif was absent in the Ftf of *S. salivarius*.

Examination of the C-terminus of the Ftf of *S. salivarius* following exonuclease deletions of the 3'OH region of the *ftf* gene and expression of the mutated alleles in *Streptococcus gordonii* Challis LGR2, showed that the C-terminus was required for cell attachment, as only the complete gene product remained cell-bound. This cell-bound activity was released in the presence of sucrose, suggesting that the mode of attachment and release of the Ftf expressed in *S. gordonii* was similar to that in *S. salivarius*. Site-directed mutated Ftfs specifically designed to examine the role of the C-terminal domains implicated both the hydrophobic and proline-glycine- and threonine-serine-rich wall-associated domain in stabilizing the Ftf on the surface of the cell.

CONTENTS

STATEMENT OF AUTHORSHIP.....	(i)
ACKNOWLEDGEMENT.....	(ii)
PREFACE.....	(iii)
SUMMARY.....	(iv)
CONTENTS.....	(v)
ABBREVIATIONS.....	(x)
1 INTRODUCTION.....	1
1.1 Microbial Fructan Synthesis.....	1
1.1.1 Oral Microbial Colonization and Dental Diseases.....	1
1.1.2 Extracellular Polysaccharide Synthesis from Sucrose in the Oral Cavity	2
1.1.3 Oral Microbial Fructans.....	2
1.1.4 Structure of the Fructans Produced by Oral Streptococci.....	3
1.1.5 Fructosyltransferases Produced by Other Genera of Bacteria.....	5
1.1.6 Purification and Properties of Non-Streptococcal Fructosyltransferases	6
1.1.7 Purification and kinetic Properties of the levansucrase (SacB) of Bacilli	8
1.1.8 Key Amino Acids in the SacB of <i>B. subtilis</i>	10
1.1.9 Purification and Properties of Streptococcal Fructosyltransferases.....	11
1.2 Attachment of Proteins to Cell Surfaces.....	13
1.2.1 Attachment of Proteins to Plasma Membranes.....	13
1.2.1.1 Non-covalent Binding of Membrane Proteins.....	14
1.2.1.2 Covalent Binding of Proteins to Membranes – Lipoproteins.....	16
1.2.2 Attachment of Proteins to Peptidoglycan.....	20
1.2.2.1 Synthesis and Structure of Peptidoglycan.....	20
1.2.2.2 Covalent Binding of Proteins to Gram-positive Cell Walls.....	22
1.2.3 The Transpeptidase (Sortase) Responsible for Processing the LPXTG Site.....	26
1.2.4 An Alternative Cell Surface Protein Binding Mechanism in Gram-positive Bacteria.....	28
1.2.5 Aims of the Present Study.....	28

2 ROUTINE METHODS.....	29
2.1 Reagents, Chemicals and Enzymes.....	29
2.2 Bacterial Strains, Bacteriophage, Plasmids and Phagemids.....	30
2.3 Media and Growth Conditions.....	33
2.4 DNA Manipulations.....	33
2.4.1 Purification of Bacteriophage- λ DNA.....	33
2.4.2 Rapid Purification of Phagemid DNA.....	34
2.4.3 Small Scale Purification of Plasmid DNA from <i>Streptococcus gordonii</i> ..	35
2.4.4 Large Scale Purification of Phagemid DNA.....	35
2.4.5 Preparation of Single-Stranded Phagemid DNA.....	36
2.4.6 Spectrophotometric Determination of Nucleic Acid Concentration.....	37
2.4.7 Electrophoresis of DNA on Agarose Gels.....	37
2.4.8 Recovery of DNA from Agarose Gels.....	37
2.4.9 Digestion of DNA with Restriction Endonucleases.....	38
2.4.10 Dephosphorylation of DNA.....	38
2.4.11 DNA Ligation.....	38
2.4.12 Exonuclease Digestion of DNA.....	38
2.4.13 Preparation of Competent <i>E. coli</i>	39
2.4.14 Transformation of <i>E. coli</i> by Electroporation.....	39
2.4.15 Natural Transformation of <i>S. gordonii</i>	40
2.4.16 DNA Sequencing.....	40
2.5 Detection of Fructosyltransferase Activities and Protein Manipulations	41
2.5.1 Detection of Sucrase Activity.....	41
2.5.2 Assay for Fructosyltransferase Activity	41
2.5.3 SDS-PAGE Gel Electrophoresis.....	41
2.5.4 Detection of Proteins on SDS-PAGE with Silver Staining.....	42
2.5.5 Detection of Fructosyltransferase Activity in SDS-PAGE.....	42
2.5.6 Electoblotting of Proteins from SDS-PAGE onto PVDF Membranes..	43
2.6 Data Processing and Bioinformatic Analyses.....	43
3 ANALYSIS OF THE REGION FLANKING THE <i>ftf</i> GENE IN	
BACTERIOPHAGE λA-39.....	44
3.1 Introduction.....	44
3.2 Methods	44
3.2.1 Sequencing of DNA.....	44

3.2.2 DNA Sequence Analysis.....	44
3.3 Results and Discussion.....	45
3.3.1 Physical Map of <i>Streptococcus salivarius</i> ATCC 25795 Chromosomal DNA in λ A-39.....	45
3.3.2 The ORF Downstream of the <i>ftf</i> Gene Codes for a <i>FURI</i> Homologue.	49
3.3.3 The ORF Downstream of the Putative <i>FURI</i> Gene Codes for a ClpP Homologue.....	49
3.3.4 DNA upstream of the <i>ftf</i> Gene Codes for an Insertion Sequence.....	52
4 THE CELL-ASSOCIATED FRUCTOSYLTRANSFERASE OF <i>STREPTOCOCCUS SALIVARIUS</i>.....	56
4.1 Introduction.....	56
4.2 Methods.....	56
4.2.1 DNA Sequencing.....	56
4.2.2 DNA Sequence Analysis.....	59
4.3 Results.....	59
4.3.1 Sequencing of the <i>ftf</i> Gene of <i>S. salivarius</i>	59
4.3.2 The <i>S. salivarius</i> Ftf Shares Common Ancestry with Equivalent Enzymes.....	59
4.3.3 The <i>S. salivarius</i> Ftf Contains a Novel Proline-rich C-terminal Region	67
4.3.4 The C-terminus of the Ftf of <i>S. salivarius</i>	69
4.3.4.1 Cloning of Deletion Mutated <i>ftf</i> Genes into the Shuttle Vector pVA838.....	69
4.3.4.2 Expression of <i>S. salivarius</i> Ftf Activity in <i>S. gordonii</i> LGR2	71
4.3.4.3 Localization of Mutated Ftf Activities in <i>S. gordonii</i>	71
4.3 Discussion.....	71
5 DEVELOPMENT OF A TECHNIQUE FOR MULTIPLE SITE- DIRECTED MUTAGENESIS OF THE <i>ftf</i> GENE CONTAINING PALINDROMIC SEQUENCES.....	75
5.1 Introduction.....	75
5.2 Methods.....	75
5.2.1 Chemicals and Enzymes.....	75
5.2.2 Bacterial Strains, Phagemids, Bacteriophage and Growth Conditions	75
5.2.3 DNA Manipulations and Mutagenesis.....	76
5.2.3.1 The T7 Polymerase Method.....	77

5.2.3.2 The Vent Polymerase Method.....	77
5.2.4 Transformation and Selection.....	78
5.3 Results.....	78
5.3.1 Formation of Deletions Using Standard Site-Directed Mutagenesis Protocols.....	78
5.3.1.1 Single-Stranded Mutagenesis.....	80
5.3.1.2 Double-Stranded Mutagenesis.....	80
5.3.2 Production and Mutagenesis of a 5.4 kbp Phagemid containing a Partial <i>ftf</i> Gene.....	80
5.3.3 Examination of the Stem-Loop Termination Points and Mutagenesis of the <i>ftf</i> Gene by Modified Techniques.....	85
5.3.4 Sequencing of Mutated <i>ftf</i> Genes.....	88
5.4 Discussion.....	88
6 ROLE OF C-TERMINAL DOMAINS IN SURFACE ATTACHMENT OF THE FRUCTOSYLTRANSFERASE OF <i>STREPTOCOCCUS SALIVARIUS</i>.....	93
6.1 Introduction.....	93
6.2 Methods.....	95
6.2.1 Bacterial Strains and Growth Conditions.....	95
6.2.2 Plasmids and Phagemids.....	95
6.2.3 Site-Directed Mutagenesis.....	95
6.2.4 Formation of Truncated <i>ftf</i> Genes.....	95
6.2.5 Detection of Ftf Activity.....	96
6.2.6 Release of Cell-bound Ftf Activity in Washed Cell Suspensions	96
6.2.7 Fractionation of <i>E. coli</i> Cells.....	96
6.2.8 Data Base Searches.....	96
6.3 Results.....	97
6.3.1 Expression of Mutated Ftf's in <i>E. coli</i>	97
6.3.2 Cellular Location of Mutated Ftf's in <i>S. gordonii</i>	97
6.4 Discussion.....	100
7 ANALYSIS OF THE SUCROSE-RELEASED FORMS OF FRUCTOSYLTRANSFERASE.....	102
7.1 Introduction.....	102
7.2 Methods.....	102

7.2.1 Bacterial Strains, Phagemids and Growth Conditions.....	102
7.2.2 Release of Cell-bound Ftf Activity in Washed Cell Suspensions	103
7.2.3 The Effect of Protease Inhibitors on the Release of Ftf from the Cell Surface.....	103
7.2.4 Preparation of Ftf for C-terminal Analysis.....	103
7.2.5 Western Blotting and N-terminal Sequence Analysis of the Ftf..	105
7.3 Results.....	105
7.3.1 Sucrose-released Forms of Ftf.....	105
7.3.2 Effect of Growth Phase on the Form of Sucrose-released Ftf.....	108
7.3.3 Effect of Exposure to Sucrose on the Form of Ftf.....	108
7.3.4 Effect of Proteinase Inhibitors on the Release of Ftf from <i>S. salivarius</i>	108
7.3.5 C-terminal Analyses.....	114
7.3.6 N-terminal Sequencing.....	114
7.4 Discussion.....	115
8 GENERAL DISCUSSION.....	119
APPENDIX I.....	137
APPENDIX II.....	138
BIBLIOGRAPHY.....	139
RESEARCH PAPERS.....	173

ABBREVIATIONS

λ	lambda phage
$^{\circ}\text{C}$	degrees Celsius
Ω	ohm
%	percentage
A_{260}	absorbancy at 260 nm
A_{280}	absorbancy at 280 nm
A_{600}	absorbancy at 600 nm
ADP	adenosine 5'-diphosphate
ANGIS	Australian National Genomic Information Service
APAF	Australian Proteome Analysis Facility
ATP	adenosine 5'-triphosphate
BHI	brain heart infusion
bp	base pair
Bq	becquerel
C-1	carbon at position 1
C-2	carbon at position 2
C-6	carbon at position 6
Ci	curie
cpm	counts per minute
C-terminus	carboxyl-terminus
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	a mixture of four types of deoxynucleotide: deoxyadenosine 5'-triphosphate (dATP), deoxyguanosine 5'-triphosphate (dGTP), deoxycytidine 5'-triphosphate (dCTP) and deoxythymidine 5'-triphosphate (dTTP)
dsDNA	double-stranded DNA
DTT	dithiothreitol
E	enzyme

EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N', -tetraacetic acid
Em ^R	erythromycin resistant
Fru	fructose
Ftf	fructosyltransferase
Ftf ⁻	fructosyltransferase negative
Ftf ⁺	fructosyltransferase positive
<i>ftf</i>	gene coding for fructosyltransferase
<i>g</i>	force of gravity
g	gram
Glc	glucose
Gtf	glucosyltransferase
h	hour
in	inch
IPTG	β -D-thiogalactoside
IS	insertion sequence
HPLC	High pressure liquid chromatography
k	kilo
K_{cat}	catalytic constant
K_{m}	Michaelis constant
L	litre
lb	pound
LB medium	Luria-Bertani medium
m	milli-
M	molar
M Ω	mega ohm
mA	milli-ampere
MALDI	matrix assisted laser desorption ionization (mass spectrometry)
mg	milli-gram
min	minute
ml	milli-litre
mM	milli-molar
mol	mole

M_r	relative molecular weight
nm	nanometre
nM	nanomolar
N-terminus	amino-terminus
ORF	open reading frame
PAGE	polyacrylmide gel electrophoresis
PAS reaction	periodic acid Schiff reaction
PEG	polyethyleneglycol
pfu	plaque forming units
pH	degree of acidity
PHYLP	Phylogeny interference package
pI	isoelectric point
PMSF	phenylmethylsulfonyl fluoride
PVDF	Polyvinylidenedifluoride
RBS	ribosome binding site
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
s	second
S1 reaction mix	nuclease S1 reaction buffer
S1 stop mix	nuclease S1 stop buffer
Sec	secretion
SDM	semi-defined medium
SDS	sodium dodecyl sulfate
SDS PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHA	saliva-coated hydroxyapatite
ssDNA	single-stranded DNA
STEP buffer	SDS, Tris, EDTA, proteinase K buffer
Suc	sucrose
SUPAMAC	Sydney University and Prince Alfred Macromolecular Analysis Centre
TBE	Tris-borate EDTA buffer
TE	Tris EDTA buffer
TEMED	N, N, N', N'-tetramethylethylenediamine

TFA	trifluoroacetic acid
THB	Todd Hewitt Broth
THBS	THB with heat inactivated horse serum
TM buffer	Tris MgSO ₄ buffer
Tris	trizma base
U	unit of enzyme activity
U- ¹⁴ C-	uniformly labelled carbon 14 isotope
UPRTase	uracil phosphoribosyltransferase
UV	ultraviolet
V	volts
v/v	volume per volume
W	watt
w/v	weight per volume
WHO	World Health Organization
X-gal	5-bromo-4-chloro-3-indolyl-galactosidase
YT medium	yeast-tryptone medium
μ	micro-
μF	micro-farad (unit for capacitance)
μl	micro-litre

CHAPTER 1

INTRODUCTION

1.1 MICROBIAL FRUCTAN SYNTHESIS

1.1.1 Oral Microbial Colonization and Dental Disease

The unique features of the oral cavity provide a habitat for the largest and most diverse population of microorganisms anywhere in the body. The temperature is maintained close to 37°C and the pH is buffered close to neutrality by saliva. Sources of nitrogen and carbon for microbial metabolism are continuously supplied by saliva in the form of salivary proteins and glycoproteins. In the healthy oral cavity, the microbial community is relatively stable, but under certain circumstances, microbial biofilms may cause major dental diseases such as dental caries and periodontal diseases. These diseases are of significant economic and social importance worldwide. It is not surprising, therefore, that the environmental factors that affect the microbial colonisation of the mouth and the mechanisms leading to the onset of the dental diseases have been the subject of considerable research.

The ability of microorganisms to adhere to surfaces is an important ecological determinant, especially in the oral cavity where swallowing and flow of saliva rapidly removes non-adherent microorganisms (Gibbons, 1984). The firmly adherent accumulation of bacterial biofilms on the teeth is known as dental plaque and this has been identified as the aetiological agent of dental caries and periodontal diseases. Due of the shearing action of saliva, dental plaque is generally restricted to the more sheltered sites such as pits and fissures found on the occlusal surfaces of the molar teeth or at the gingival margin.

Dental caries is a localized destruction of tooth enamel or cementum and if allowed to develop will progress to involve the underlying dentine. Its onset requires a susceptible tooth surface, bacteria, a fermentable carbohydrate and time (Parfitt, 1956). It has been widely accepted that dental caries can be attributed to acid production by oral microorganisms found in dental plaque (Hogg, 1990; Jacques, 1993). These acids may be trapped at the tooth surface by the overlying dental plaque for sufficient time to affect the solubility of the hydroxyapatite crystals of the tooth enamel or cementum. Although the use of various fluorides in the water supply, mouthwashes and toothpastes have greatly reduced the incidence of dental caries, it remains a serious problem in many countries (WHO, 1986, 1994).

1.1.2 Extracellular Polysaccharide Synthesis from Sucrose in the Oral Cavity

The nutrients required for oral microbial metabolism are continuously supplied by saliva and intermittently from exogenous foodstuffs in the diet. Sucrose is a major dietary carbon source for humans in many western countries (Department of Health, 1989; Glinsmann *et al.*, 1986) and studies have shown that sucrose is one of the major dietary agents in the aetiology of dental caries (Rugg-Gunn, 1989).

Based on animal studies and cross-sectional human studies, it is clear that oral streptococci are associated with dental caries (Loesche, 1986). Oral streptococci such as the mutans streptococci and *Streptococcus salivarius* produce a wide variety of sucrose-metabolizing enzymes, which appear to be responsible for the critical role of sucrose in cariogenesis. Fructosyltransferases (Ftfs) and glucosyltransferases (Gtfs) that synthesize extracellular fructan and glucan polymers respectively are two of the major categories of sucrose-metabolizing enzymes. These extracellular enzymes, existing either free or bound to the cell surface, can utilize the relatively high free energy of hydrolysis of sucrose to form their respective polysaccharides without the need for the input of energy in the form of ATP or other high energy intermediates (Jacques, 1993).

1.1.3 Oral Microbial Fructans

Fructan was the first polysaccharide identified in dental plaque (Niven *et al.*, 1941). The large M_r ($>10^6$) and solution viscosity of fructan (Birkhed *et al.*, 1979; Ehrlich *et al.*, 1975) make it hard for fructan to diffuse out of the dental plaque matrix (Manly and Richardson, 1968). The role of fructans in the aetiology of dental caries is often considered in conjunction with extracellular glucans (Hardie, 1986; Gehring, 1981; Montville *et al.*, 1978). For example, Gibbons and Nygaard (1968) surveyed a number of cariogenic and non-cariogenic oral streptococci and found that Ftfs were present along with Gtfs in the culture fluid of more than one species of streptococci. While the formation of water-insoluble glucans has been shown to be a critical factor in dental plaque formation by cross-linking bacteria and by interacting with receptors within the acquired pellicle to provide a mechanism by which the dental plaque biofilms adheres to the tooth surface (Hamada and Slade, 1980; Gibbons, 1983), fructans have been found to enhance the formation of glucan-based bacterial biofilms on tooth surfaces and serve as a reserve of extracellular carbohydrate (Gibbons, 1968).

Even though the amount of fructan represents only a small percentage (~5%) of the total polysaccharides in dental plaque, its rate of synthesis is higher than that of glucan in both plaque and saliva (Wood, 1967; Rölla *et al.*, 1983). The rapid accumulation of fructans in

dental plaque following consumption of sucrose has been demonstrated with human volunteers (Gold *et al.*, 1974). The low level of fructan in dental plaque is due to its rapid hydrolysis by bacterial fructanases (Wood, 1967; Manly and Richardson, 1968; Walker, *et al.*, 1983; Jacques, *et al.*, 1985; Takahashi, *et al.*, 1985; Burne *et al.*, 1987, 1996). It has been shown that following the exhaustion of dietary carbohydrate, the expression of fructanases could be up-regulated by fructan (Burne *et al.*, 1995). The fructose released from the hydrolysis of fructan is transported by the phosphoenolpyruvate-dependent phosphotransferase systems (Gauthier *et al.*, 1984) into the bacterial cell and subsequently catabolized to lactic acid or other acids, giving rise to prolonged periods of low pH in dental plaque and the subsequent demineralization of tooth enamel. (Loesche, 1986; Jacques, 1993).

1.1.4 Structure of the Fructans Produced by Oral Streptococci

Fructans produced from sucrose in the oral cavity were originally ascribed to *S. salivarius* (Niven *et al.*, 1941; Wood, 1967), a *Streptococcus* first described by Andrewes and Horder in 1906. It has been shown that *S. salivarius* can be isolated in greater numbers from dental plaque on teeth when sucrose consumption increases (Carlsson, 1968) even though this species adheres poorly to tooth surfaces and is most often found in high numbers on epithelial tissues, especially the dorsum of the tongue (Gibbons and van Houte, 1975). Recent refined taxonomic studies have confirmed that *S. salivarius* is the dominant *Streptococcus* on the dorsum of the tongue (Milnes *et al.*, 1993a) but that the numbers of this species persisting in dental plaque is larger than that previously believed (Milnes *et al.*, 1993b). *S. salivarius* is non-haemolytic and can be readily identified on sucrose-containing solid growth medium by its appearance as large mucoid colonies due to the production of extracellular fructan (Niven *et al.*, 1941). *S. salivarius* has been shown to produce fissure and root surface caries in mono-infected rats (Kelstrup and Gibbons, 1970) and may well do so in humans, which is a problem of growing concern with an aging population (Ravald, 1994). *S. salivarius* has been classified into the high-carries category, along with the members of the mutans streptococci (Drucker *et al.*, 1984) and attempts have been made to study the fructans produced by this species in order to uncover the mechanism of its pathogenicity (Wood, 1967; Ebisu *et al.*, 1975; Stivala *et al.*, 1975; Seymour *et al.*, 1979a, b; Marshall and Weigel, 1980a, b; Garszczynski and Edwards, 1973; Jacques and Wittenberger, 1981; Jacques, 1984; Pitty and Jacques, 1987; Pitty *et al.*, 1989).

The structure of the fructan produced by *S. salivarius* has been characterized as a levan in which the D-fructofuranose units are joined by a β -(2 $\rightarrow$ 6) linkage with β -(2 $\rightarrow$ 1) branch points. This contrasts with the other major bacterial fructan, inulin, formed by the mutans

streptococci in which the D-fructofuranose units are joined with β -(2 $\rightarrow$ 1) fructosyl-fructose linkages with β -(2 $\rightarrow$ 1) branch points. The average repeating-unit of levans is 8-11 β -D-fructofuranose residues (Ebisu *et al.*, 1975; Stivala *et al.*, 1975; Seymour *et al.*, 1979a, b; Marshall and Weigel, 1980a, b). The linkage between the fructosyl unit in levan is shown in Figure 1.1. The branch point β -(2 $\rightarrow$ 1)-bonds of levan are readily broken by acid hydrolysis prior to hydrolysis of the main chains (Lauren *et al.*, 1975).

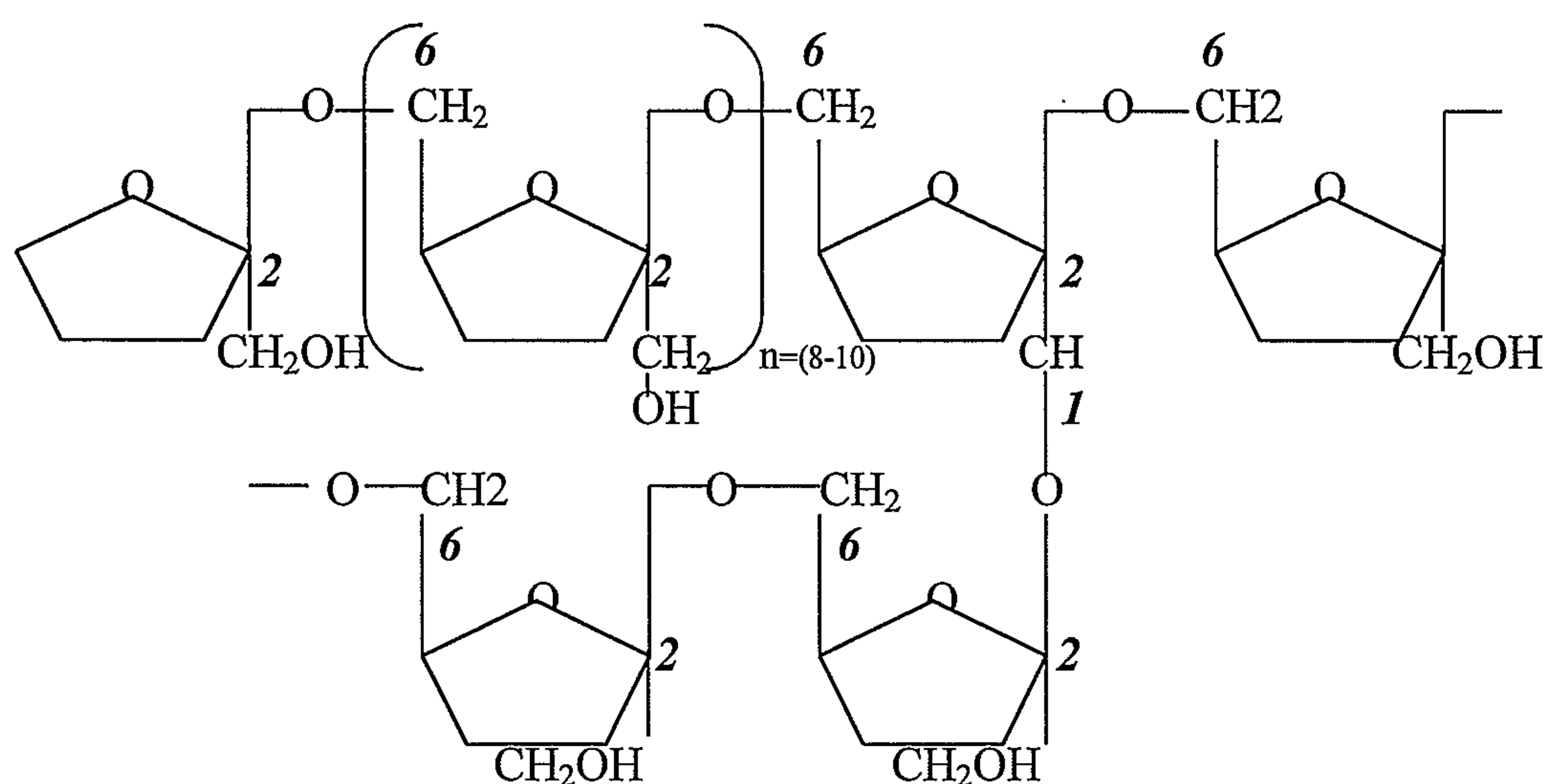


Figure 1.1 The levan structure of the fructan produced by *S. salivarius* (after Song, 1999)

The levans elaborated by *S. salivarius* have a high M_r of $5-30 \times 10^6$ and possess multiple-branched chains as shown in Figure 1.2. As such the native levans are arborescent, highly branched, near-spherical structures having greatest density at their inner centre and lowest towards their periphery. These levans are water-soluble and behave hydrodynamically as compact particles with spherical symmetry, a property that results in their relatively low intrinsic viscosity (Stivala *et al.*, 1975).

The highly branched and spherical structure of the levan produced by *S. salivarius* allows it to be retained in dental plaque as a poorly-diffusible carbohydrate reserve while being effectively hydrolyzed by exo-hydrolases (fructanases) (Da Costa and Gibbons, 1968; Ehrlich *et al.*, 1975; Marshall and Weigel, 1980b). Hydrolysis probably occurs at the more accessible outer branch points and proceeds along each chain towards the centre core (Enrlich *et al.*, 1975; Marshall and Weigel, 1980 a, b; Stivala and Zweig, 1981; Khorramian and Stivala, 1982).

Fructans are also synthesized from sucrose in the oral cavity by the mutans streptococci

inducible levansucrase (SacB) from *Bacillus subtilis* and a constitutive levansucrase-containing extract from a strain of *Aerobacter*. Since then, Ftf's have been identified in a wide variety of microorganisms including the Gram-negative bacteria, *Acetobacter diazotrophicus* (Hernandez *et al.*, 1995), *Acetobacter levanicum* (Hestrin *et al.*, 1943; Hestrin and Avineri-shapiro, 1944), *Erwinia amylovora* (Gross *et al.*, 1990), *Erwinia herbicola* (Cote and Imam, 1989), *Gluconobacter oxydan* (Elisashivili, 1980), *Pseudomonas syringae* (Hettwer *et al.*, 1995), *Rhanella aquatilis* (Ohtsuka *et al.*, 1992), *Zymomonas mobilis* (Lyness and Doelle, 1983) and the Gram-positive bacteria, *B. subtilis* (Hestrin *et al.*, 1943; Dedonder, 1966; Gonzy-Treboul *et al.*, 1975), *B. amyloliquefaciens* (Mantsala and Puntala, 1982), *Bacillus natto* (Takahama *et al.*, 1991), *Bacillus polymyxa* (Hestrin *et al.*, 1943; Han, 1989) *Bacillus stearothermophilus* (Li *et al.*, 1997), *S. salivarius* (Hehre, 1945; Jacques, 1984, 1993), *S. mutans* (Sato *et al.*, 1984), *S. rattus* (Baird *et al.*, 1973; Ebisu *et al.*, 1975), *S. cricetus* (Birkhed *et al.*, 1979) and *Actinomyces naeslundii* (*viscosus*) (van der Hoeven *et al.*, 1976; Pabst, 1977; Pabst *et al.*, 1979).

The genes coding for Ftf activity have been cloned and sequenced from a number of these bacteria. They include *S. mutans* (Shiroza and Kuramitsu, 1988), *B. subtilis* (Steinmetz *et al.*, 1985), *B. amyloliquefaciens* (Tang *et al.*, 1990), *E. amylovora* (Geier and Geider, 1993), *Z. mobilis* (Song *et al.*, 1993; Kyono *et al.*, 1995), *A. diazotrophicus* (Arrieta *et al.*, 1996), *B. stearothermophilus* (Li *et al.*, 1997), *P. syringae* (Hettwer *et al.*, 1998) and most recently, *A. naeslundii* (Bergeron *et al.*, 2000).

1.1.6 Purification and Properties of Non-Streptococcal Fructosyltransferases

Levansucrases of *A. levanicum* (*E. herbicola*) have been isolated from three different sources; a cell-free extract of an autolysed culture (Hestrin *et al.*, 1943), homogenates of cells (Ebert and Schenk, 1968) and the media containing the secreted enzyme (Cote and Imam, 1989). The resulting enzymes differ largely in M_r and the K_m value for sucrose, probably due to the different techniques used to obtain them (Cote and Ahlgren, 1993).

A. naeslundii is commonly associated with dental plaque and produces a constitutive and extracellular levansucrase (Snyder *et al.*, 1967; van der Hoeven *et al.*, 1976; Pabst, 1977; Pabst *et al.*, 1979). A levansucrase has been purified 165-fold from the culture supernatant of *A. naeslundii* Ny1 by DEAE-anion exchange chromatography and isoelectric (van der Hoeven *et al.*, 1976). The purified enzyme exhibits maximal activity at pH 6.8 and 45°C. The metal ions, Ca^{2+} and Mg^{2+} were found to activate the enzyme, whereas Hg^{2+} and Cu^{2+} inhibit it.

Pabst (1977) also purified a levansucrase from the culture supernatant of *A. naeslundii* T-14V following ammonium sulfate precipitation, DEAE-cellulose chromatography, gel

filtration and two hydroxyapatite chromatographic steps with a 37 000-fold increase in specific activity. The levansucrase possessed a native M_r of 220 000 according to gel filtration chromatography and consisted of subunits of M_r 80 000 as determined by SDS-PAGE (Pabst, 1977). In contrast to the levansucrase purified by van der Hoeven *et al.* (1976), maximal activity was observed at pH 6.0 and 37°C and no activation was observed with the addition of Ca^{2+} or Mg^{2+} . The K_m for sucrose was estimated to be 12 mM (Pabst, 1977). The fructan product was a branched levan of a high M_r of 10^8 (Pabst, 1977). A later study discovered that *A. naeslundii* produced both a soluble extracellular and a cell wall-associated levansucrase. The solubilized cell wall-associated form released from the cell surface by lysozyme digestion appeared to be identical to the native soluble extracellular form (Pabst *et al.*, 1979). It has also been shown that the levansucrase of *A. naeslundii* remains tightly bound to its levan product (Pabst, 1977) but not to exogenously added levan (Pabst *et al.*, 1979). It has not been determined whether the levansucrase purified by van der Hoeven *et al.* (1976) and that by Pabst (1977) are inherently different enzymes or that the differences in their enzymic properties are due to experimental error.

The plant pathogenic bacterium, *P. syringae*, produces a levan as a protective spacer between the plant cell wall and the bacteria in order to inhibit the recognition and initiation of a hypersensitive response by the host plant (Hettwer *et al.*, 1995). The extracellular levansucrase of *P. syringae* pv. Phaseolicola has been purified from the cell supernatant by gel chromatography. The purified enzyme shows a M_r of 45 000 on SDS-PAGE and M_r of 68 000 by gel filtration. Optimal activity is observed at pH 5.8 to 6.0 and 60°C. The enzyme shows high tolerance to denaturing agents, proteinases, and repeated freezing and thawing (Hettwer *et al.*, 1995).

A. diazotrophicus, a nitrogen-fixing bacterium associated with sugar cane, secretes a levansucrase which is expressed constitutively and constitutes more than 70% of the total protein secreted by strain SRT4 (Hernandez *et al.*, 1995). The purified enzyme has a M_r of 58 000 and pI of 5.5. The enzyme shows optimal activity at pH 5.0 and is inhibited by Hg^{2+} . Addition of Ca^{2+} or the metal chelator EDTA does not affect enzyme activity (Hernandez *et al.*, 1995).

An extracellular levansucrase is also secreted by *Z. mobilis*, a fermentative bacterium that efficiently produces ethanol from glucose. Since the level of levansucrase expression by this organism is very low even when induced by sucrose, the purification of this enzyme has not been reported. However, the gene that codes for this enzyme has been cloned and sequenced (Song *et al.*, 1993; Kyono *et al.*, 1995).

1.1.7 Purification and Kinetic Properties of the Levansucrase (SacB) of Bacilli

SacBs of bacilli are the most extensively studied Ftf's. An extracellular sucrose-induced SacB was first purified from *B. subtilis* BS5 in Dedonder's laboratory (Dedonder, 1966). The enzyme shows optimal activity at pH 6.0 and a K_m for sucrose of 27 mM. Addition of the metal ions, Fe^{3+} , Al^{3+} , or Zn^{2+} increased the heat stability of the enzyme (Dedonder, 1966). The average M_r determined by three different methods, sedimentation equilibrium, SDS-gel electrophoresis and gel filtration, was 54 000 (Gonzyl-Treboul *et al.*, 1975).

Several mutants of *B. subtilis* have been isolated, in which the extracellular SacB is expressed constitutively (Lepesant and Dedonder, 1969). This made it possible to prepare a large quantity of purified enzyme with relative ease. An X-ray crystallographic analysis of the SacB purified from the constitutive mutant *B. subtilis* BS5C4 has been determined and a model of the enzyme constructed to 3.8Å resolution (LeBrun and van Rapenbusch, 1980). The molecule was shown to be a elongated prolate ellipsoid, with the dimensions 26 x 32 x 117 Å and a narrow region near the centre, which may allow some degree of flexibility around the active site region.

Detailed kinetic studies have been performed with the SacB of *B. subtilis* and the mechanism of catalysis elucidated. The transfructosylation reaction catalyzed by the enzyme is a multiple-reaction in which sucrose or raffinose acts as the fructosyl donor and substances possessing alcoholic (hydroxyl) groups such as water, sucrose, glucose or oligolevan can serve as the fructosyl acceptor as shown below (Dedonder, 1966):



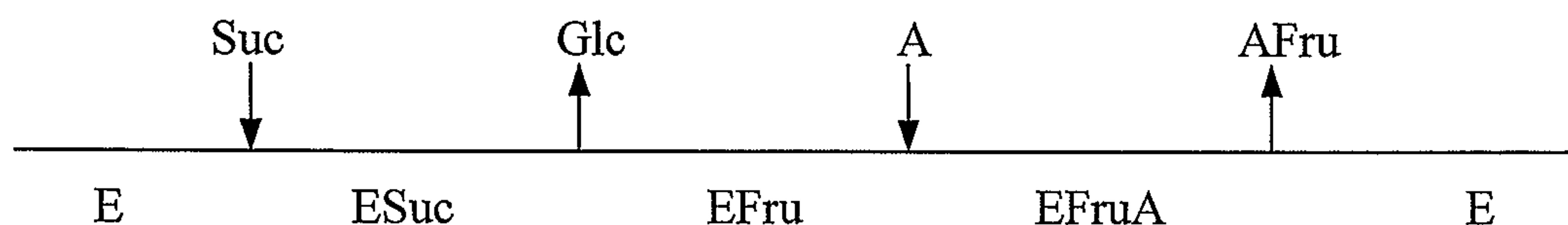
SacB can hydrolyze small levans, but not inulin and the hydrolysis is stopped at branch points. An exchange reaction is also observed when SacB is mixed with sucrose and [U- ^{14}C]-labelled glucose (Dedonder, 1966), such that:



Based on these findings, it was initially concluded that SacB transfers a fructosyl residue from sucrose to the primary alcohol at C-6 of the fructosyl residue at the non-reducing end of a growing levan chain possibly via a fructosyl-enzyme intermediate complex (Dedonder, 1966).

The mechanism of the reaction catalyzed by the SacB of *B. subtilis* was further investigated by Chambert *et al.* (1974) by careful analysis of the initial velocities of the

fructosyl transfer reactions. The results strongly supported a Ping Pong mechanism of catalysis involving the participation of a fructosyl-enzyme intermediate. Thus the enzyme reaction can be described using the short-hand notation of Cleland (Cleland, 1973) as follows:



where E, Suc, Glc, Fru and A represent the enzyme, glucose, fructose and acceptor respectively (Chambert and Gonzyl-Treboul, 1976a). A multiple chain elongation mechanism has been proposed for the extension of the levan chain. It has been argued that the fructosyl residues are added randomly to all molecules of levan acceptor rather than successively to the same molecule of the levan (Chambert *et al.*, 1974).

The Ping Pong mechanism was further confirmed by thermodynamic and kinetic studies (Chambert and Gonzyl-Treboul, 1976b). A simplified method to evaluate the kinetic constants of the exchange and hydrolytic reaction was developed and the resulting data (Chambert and Gonzyl-Treboul, 1976b) was in close agreement with those predicted by a theoretical approach (Chambert *et al.*, 1974).

The mechanism by which levan synthesis is catalyzed by the SacB of *B. subtilis* was also investigated by Yamamoto *et al.* (1985) using raffinose as the substrate. The kinetic and chemical results indicated that the fructofuranoside chain of levan was extended by accepting a fructosyl residue at C-6 of the terminal fructofuranosyl residue by forming the linkage between C-2 and C-6. It was proposed that the enzyme strongly bound to the growing end of the levan chain and the fructosyl residues were added one at a time onto the levan (Yamamoto *et al.*, 1985).

The proposed fructosyl-enzyme intermediate in the Ping Pong reaction has been isolated from the reaction mixture of the enzyme and sucrose (Chambert and Gonzyl-Treboul 1976a). Proteolytic hydrolysis of the trapped complex resulted in the identification of an aspartic acid residue as the amino acid to which the fructosyl residue attached. The lability of the fructosyl-aspartate bond under mild alkaline conditions suggested that the fructosyl residue was linked through an ester bond involving the β -carboxyl of the aspartic acid residue (Chambert and Gonzyl-Treboul, 1976a).

The only other Ftf whose kinetic properties have been studied in detailed prior to the cloning of the *ftf* gene of *S. salivarius* (Song and Jacques, 1999a) is the levansucrase secreted by *A. diazotrophicus*. This enzyme catalyses fructosyl transfer from sucrose by a Ping Pong

mechanism involving the formation of a transient fructosyl-enzyme intermediate in a similar manner to the levansucrase of *B. subtilis* (Hernandez *et al.*, 1995). Both enzymes can synthesize fructan from sucrose alone when the sucrose concentration ≥ 50 mM. Most of the kinetic parameters of the two enzymes are of the same magnitude, except that the levansucrase of *A. diazotrophicus* produces large amounts of 1-kestotriose and kestotetraose from sucrose, while the SacB of *B. subtilis* can synthesize fructan with a degree of polymerization >15 without accumulation of small oligofructans (Hernandez *et al.*, 1995).

1.1.8 Key Amino Acids in the SacB of *B. subtilis*

A number of key amino acids have been identified in the SacB of *B. subtilis*. For example, Chambert and Petit-Glatron (1989) discovered that increasing the concentration of organic solvent, such as 1,4-dioxane, acetone or acetonitrile, in the reaction mixture of SacB of *B. subtilis* resulted in an increase in fructan yield. In the presence of 60% acetonitrile, the hydrolase activity of the enzyme was completely inhibited and the yield of fructan reached 100%. The K_m values for sucrose and raffinose were unaffected but the K_{cat} values were increased 5-fold (Chambert and Petit-Glatron, 1989).

Following the observation that the SacB of *B. subtilis* preferentially maintains its polymerase activity in the presence of high concentrations of organic solvents, Chambert and Petit-Glatron (1991) demonstrated that the polymerase and hydrolase activities of the enzyme could be separately modulated by site-directed mutagenesis. The mutation R331H resulted in an enzyme of much lower polymerase activity. Substitution of R³³¹ with lysine, serine or leucine led to a loss of the synthesise of levan from sucrose alone. The mutated enzymes could only catalyze the formation of the trisaccharide, kestose (Chambert and Petit-Glatron, 1991).

The secretion of SacB by *B. subtilis* has been characterized as a two-step process. The first step involves the proteolytic cleavage of the N-terminal signal sequence to give a mature membrane-associated form and the second, rate limiting step, involves the secretion and concomitant conformational change of the resulting mature form (Petit-Glatron *et al.*, 1987). The second step of the secretion has been found to be coupled to the continuous uptake of Fe^{3+} by the bacteria (Chambert and Petit-Glatron, 1988). Studies on the *in vitro* refolding of the levansucrase following its denaturation by urea revealed that the rate of refolding was greatly enhanced by Fe^{3+} whereas the ion chelator, EDTA, prevented correct folding (Chambert *et al.*, 1990). Like Fe^{3+} , addition of Ca^{2+} ions has also been shown to promote the refolding of SacB at physiological temperature and pH. In the absence of either of these ions, the SacB remains in a predominantly unfolded form (Chambert and Petit-Glatron, 1990). It

has been proposed that the energy required for the vectoral diffusion of the protein from the membrane phase to the external aqueous phase is provided by the energy change accompanying the folding process and the efficiency of such a coupling could be strongly dependent on the rate of the folding process (Chambert *et al.*, 1990). Consequently changes in the external concentration of any effector, such as Fe^{3+} or Ca^{2+} , which modulates protein folding, also affects the efficiency of the second step in the secretion of the active form of the enzyme. Mutational analysis of the SacB has shown that less of the mutated enzymes, G366D and G336V are secreted by *B. subtilis* (Petit-Glatron *et al.*, 1990). The refolding process mediated by Fe^{3+} or Ca^{2+} is not observed in these mutated enzymes. However, the mutation G366S does not affect the secretion or refolding properties of the enzyme. It was therefore proposed that two requirements are essential for the secretion of the enzyme. The first is structural flexibility of the protein, as indicated by the important role of G³⁶⁶ in the folding process, and the second, is structural stability under extracellular conditions, as highlighted by the involvement of the metal ions, Fe^{3+} or Ca^{2+} as essential cofactors in coupling the folding and translocation processes (Petit-Glatron *et al.*, 1990).

Another highly conserved region is the so called 'sucrose box', SGSA----D----LYYT, which is also found in the Ftf of *S. mutans* (Sato and Kuramitsu, 1988). The 'sucrose box' has also been found in other enzymes catalyzing the transfer of fructose from sucrose such as sucrose-6-phosphate hydrolase (Sato and Kuramitsu, 1988) which otherwise show little homology to the Ftfs. However, the actual role of the 'sucrose box' structure in catalysis was not determined for any Ftf or sucrose hydrolyzing enzyme prior to the sequencing of the *ftf* gene of *S. salivarius* (Chapter 4; Song and Jacques 1999b).

1.1.9 Purification and Properties of Streptococcal Fructosyltransferases

An extracellular Ftf is produced constitutively by *S. mutans* JC2 when grown on glucose (Carlsson, 1970). The enzyme has been extracted from culture supernatants *in vitro* by hydroxyapatite adsorption and purified 200-fold to give a 10% overall yield following hydroxyapatite chromatography and isoelectric focusing. The Ftf exhibits optimal activity around pH 6 and at 40°C, and has a pI of 4.2 (Carlsson, 1970). The enzyme is inhibited by EDTA and activated by Ca^{2+} ions. The Ftf activity is completely abolished in the presence of 0.1 mM Hg^{2+} and is reduced to 40% by 1 mM Pb^{2+} , Ag^+ , Ni^{2+} or Cu^{2+} ions (Carlsson, 1970). The high M_r fructan produced has been described as an inulin type with 5% β -(2→6) branch points (Rosell and Birkhed, 1974).

The constitutive extracellular Ftf produced by *S. mutans* FA1 was co-purified with a Gtf at each step of ammonium sulfate precipitation, gel filtration and DEAE-cellulose

chromatography (Scales *et al.*, 1975). The two enzymes were finally separated after addition of the nonionic detergent Tween 80 to break the hydrophobic interaction that bound the complex together (Figures and Edwards, 1979). Two forms of Ftf were identified with pIs of 3.0 and 1.7 respectively. Both forms showed maximum activity at around pH 6.0 and a K_m value for sucrose of 55 mM.

The *ftf* gene coding for the Ftf of *S. mutans* GS5 has been isolated from a plasmid library of *S. mutans* chromosomal DNA (Sato and Kuramitsu, 1986) and sequenced (Shiroza and Kuramitsu, 1988). The recombinant enzyme expressed by plasmid pSS22 in *Escherichia coli* was purified to near homogeneity after ammonium sulfate precipitation, gel filtration, anion-exchange chromatography and preparative SDS-PAGE. The purified Ftf exhibited one protein band of 63 000 after SDS-PAGE and was capable of synthesizing inulin-like fructan without the need of an oligofructan primer. Storage of the purified enzyme resulted in its degradation to active lower M_r species as small as 59 000.

Sequence analysis of the chromosomal insert from *S. mutans* GS5 in pSS22 revealed two small open reading frames (ORF1 and ORF2) upstream of the *ftf* gene and one downstream open reading frame (ORF3) which was transcribed in the opposite direction to that of the *ftf* gene (Shiroza and Kuramitsu, 1988). The intact *ftf* gene is AT-rich and codes for a 797-amino-acid protein with a predicted M_r of 87 600. The *ftf* gene is preceded by an inverted repeat region similar to that found upstream of the *sacB* gene of *B. subtilis*. In *B. subtilis* the inverted repeat region is considered to be involved in the regulation of SacB expression (Steinmetz *et al.*, 1985). Another inverted repeat was also found downstream of the *S. mutans ftf* gene. It has been suggested that this stem-loop structure might act as a transcription terminator for the *ftf* gene and also for the downstream ORF3 (Shiroza and Kuramitsu, 1988).

A typical Gram-positive signal sequence containing a basic region followed by a hydrophobic region was found in the N-terminus of the deduced amino acid sequence of the *ftf* gene of *S. mutans*. The deduced amino acid sequence of the N-terminus and the central region of the *ftf* gene showed high homology to the corresponding domains of the SacB of *B. subtilis* (Shiroza and Kuramitsu, 1988). While the Ftf of *S. mutans* synthesized fructans with inulin linkages, the levansucrase of *B. subtilis* produced levans. The molecular basis for this difference has not been elucidated.

The polypeptide encoded by ORF1 of *S. mutans* is of unknown function. However, ORF2 and ORF3 were initially thought to code for regulatory proteins involved in Ftf expression (Shiroza and Kuramitsu, 1988). A later study, confirmed that ORF3 encoded a *ftf* regulatory protein that bound specifically to the upstream region of the *ftf* promoter containing

the inverted repeat structure (Shibata and Kuramitsu, 1996). Binding was not observed when the inverted repeat structure was removed. It was therefore suggested that the gene product of ORF 3 might regulate Ftf expression by interacting with the inverted repeat region upstream of the *ftf* promoter (Shibata and Kuramitsu, 1996).

In contrast to *S. mutans*, the Ftf's of *S. salivarius* are usually cell-associated. The cell-bound Ftf of *S. salivarius* SS2 can be released by extracting the cell membrane fragments with 8 M LiCl (Garszczynski and Edwards, 1973). The enzyme was purified 23-fold by gel-filtration and DEAE-cellulose chromatography with a 60% overall yield. The final product, free of Gtf activity, exhibited optimal activity at pH 5.6. The Ftf possessed a M_r of 34 500 and a pI of 5.2. EDTA inhibited enzyme activity and the inhibition could be reversed by the addition of Mg^{2+} . The apparent K_m was 63 mM for Mg^{2+} and 17 mM for sucrose.

The Ftf of *S. salivarius* ATCC 25975 is also cell-bound (Jacques and Wittenberger, 1981) and early on it was suggested that it might be associated with the cytoplasmic membrane as its rate of expression was shown to be related to the rate of membrane lipid synthesis (Pitty and Jacques, 1987). The cell-associated Ftf was shown to be subject to proteolysis following its inhibition by free radicals generated by NADH oxidase (Jacques and Wittenberger, 1981; Abbe *et al.*, 1986). The addition of histidine, cysteine or Ca^{2+} prevented this proteolytic inactivation (Jacques and Wittenberger, 1981). Subsequent studies suggested that Ca^{2+} was a necessary cofactor for the cell-bound Ftf with K_m values for Ca^{2+} and sucrose of $18 \pm 2 \mu M$ and $12.0 \pm 0.1 mM$ respectively being determined for the cell-associated enzyme (Jacques, 1984). Later studies revealed that in the presence of sucrose, the cell-associated Ftf was released from the cell surface and secreted into the culture medium. However, there had been no reports of the isolation and purification of the Ftf of *S. salivarius* ATCC 25975 prior to the cloning (Pitty *et al.*, 1989) and sequencing of the *ftf* gene (Chapter 4) and subsequent kinetic analysis of the recombinant enzyme (Song and Jacques, 1999a, b). As the Ftf of *S. salivarius* is initially expressed as a 95% cell-associated enzyme, this raises the question as to how the Ftf is anchored to the cell surface and by what mechanism its substrate, sucrose, initiates its release.

1.2 ATTACHMENT OF PROTEINS TO CELL SURFACES

1.2.1 Attachment of Proteins to Plasma Membranes

The cytoplasmic membrane and cell wall of bacterial cell envelopes provide potential sites for protein attachment. Such proteins may be covalently or non-covalently associated with these structures.

1.2.1.1 Non-covalent Binding of Membrane Proteins

The cytoplasmic membrane is a bilayer of phospholipids composed of polar head groups such as phosphatidylethanolamine and phosphatidylglycerol and hydrophobic saturated and mono-unsaturated hydrocarbon tails. Many proteins are membrane-bound by virtue of the hydrophobic nature of their amino acids which insert spontaneously or via translocation (Sec)-dependent machinery into the lipid bilayer (von Heijne, 1994; von Heijne, 1997). The best-studied example of a Sec-independent protein insertion is that of the coat protein of the filamentous phage M13 (Geller and Wickner, 1985; Kuhn, 1995; Soekarjo *et al.*, 1996). The M13 coat protein is a small protein with a 23 amino acid N-terminal signal sequence (an 8 residue charged tail and 15 residue hydrophobic domain), a 20 amino acid periplasmic segment and a 19 amino acid hydrophobic domain followed by a charged tail of 11 residues (Kuhn and Troschel, 1992). The procoat initially binds electrostatically to the negatively charged cytoplasmic surface of the inner leaflet of the cytoplasmic membrane via the C-terminal positively charged residues (Gallusser and Kuhn, 1990). The hydrophobic signal peptide and transmembrane hydrophobic segment then insert spontaneously into the membrane in a U-like configuration (helical hairpin).

The hydrophobic domain of the M13 coat protein does not contain topological information, but is composed of hydrophobic residues which meet a hydrophobicity threshold required to maintain stable transmembrane integration. Many studies have focused on the length and composition of similar transmembrane domains. For example, analysis of the 23 amino acid hydrophobic anchor of the coliphage f1 gene III protein (pIII) in *E. coli* involved the sequential removal of these residues (Davis *et al.*, 1985). Removal of the 23 hydrophobic core resulted in the conversion from an integral membrane protein to a secreted periplasmic form (Davis *et al.*, 1985). Segments containing various copies of the hydrophobic repeat, Leu-Ala-Leu-Val, were then introduced into the mutant at the deletion site of the 23 amino acid hydrophobic domain. It was found that an artificial domain of 16 or more hydrophobic residues is required to anchor pIII to the cytoplasmic membrane (Davis and Model, 1985).

The effect of specific amino acids on the hydrophobicity threshold have been analyzed in both microsomes and *E. coli*. Hydrophobic stretches composed of alanines and leucines have been introduced into the central portion of interleukin 2 and the modified proteins expressed in an *in vitro* transcription-translation system in the presence of dog pancreas microsomes (Kuroiwa *et al.*, 1991). Stable membrane anchoring required 19 alanine residues, but only 9 leucines. This was comparable to studies of stop-transfer sequences in *E. coli* where various artificial hydrophobic segments were introduced into alkaline phosphatase. The minimum hydrophobicity of a 21 residue stretch capable of anchoring alkaline

phosphatase to the *E. coli* inner membrane was found to be 16 alanines and 5 leucines (Chen and Kendall, 1995). As in microsomes, it was also determined that a shorter stretch of 11 leucines was also sufficient to stop the translocation of the protein across the cytoplasmic membrane.

Statistical studies of membrane proteins from bacteria to eukaryotes (von Heijne and Gavel, 1988) first indicated the universal correlation between the topology of membrane proteins and the distribution of the positively charged amino acids, arginine and lysine in the regions flanking hydrophobic transmembrane segments (Boyd and Beckwith, 1990). According to the ‘positive-inside rule’, hydrophobic sequences orient themselves in the membrane with the positively charged tail in the cytoplasm (von Heijne, 1986; von Heijne, 1992). Not only can the orientation of membrane proteins be inverted by redistribution of net charge around the hydrophobic domain (von Heijne, 1989b), but the presence of positively charged residues also act to enhance the stop-transfer function of an artificially introduced hydrophobic domain (Kuroiwa *et al.*, 1991). It has long been speculated that the stability offered by the positively charged arginine and lysine residues is a result of the electrostatic interaction of these amino acids with the negatively charged head groups of the membrane phospholipids (Inouye and Halegoua, 1980). More recent studies provide a deeper understanding of the exact molecular interactions that control the positioning of transmembrane helices. Structurally, both arginine and lysine have long aliphatic side chains with positively charged groups at their ends. The arginine has a bulky guanido group while lysine has a terminal amino group. Studies have shown that the aliphatic stretch between carbon-1 and charged moieties are partitioned into the hydrophobic core of the membrane (Stopar *et al.*, 1996), while the basic guanido and amino groups interact with the acidic phospholipid head groups (Segrest *et al.*, 1990; Mishra *et al.*, 1994). This behavior is described as ‘snorkelling’ (Segrest *et al.*, 1990). The effects of arginine and lysine on the position of a transmembrane helix in the membrane under *in vivo*-like conditions has recently allowed the location of individual amino acids in a biological membrane to be studied (Monné *et al.*, 1998). The evidence indicates that arginine and lysine may be located either far into the transmembrane segment where the long hydrophobic side chains fully extend to the aqueous phase or they are orientated closer to the surface where the side chains are able to flex and ‘snorkel’ into the hydrophobic core of the membrane as previously suggested (Killan and von Heijne, 2000).

While it is generally accepted that the overall hydrophobicity of a peptide segment of a protein provides the primary driving force for the integration of transmembrane helices into the membrane and that the positively charged residues, arginine and lysine, affect the position of these

helices, other amino acids have been found to play a key role in positioning proteins within membranes. The role of these amino acids follows from recent advances in our understanding of the interfacial environment. The interfacial region is chemically heterogeneous composed as it is of phospholipid headgroups, water and acyl chain methylenes which have spilled into the interface from the hydrocarbon core. Joint refinement of X-ray and neutron diffraction data indicate that there is no distinct boundary between the upper acyl chain and headgroups of phospholipids (Wiener and White, 1992; Yau *et al.*, 1998). Statistical analysis of the three dimensional structures of known membrane proteins show a clear preference for tryptophan and tyrosine residues at the interfacial region (Landolt-Marticorena *et al.*, 1993; Reithmeier, 1995; Preusch *et al.*, 1998). Structurally tryptophan has a large hydrophobic aromatic ring attached to an amide group and an indole side chain. Tyrosine has similar properties with a smaller ring structure. The propensity of tryptophan and tyrosine to be found at the interfacial region implies a specific role for these amino acids in the correct folding and assembly of membrane-bound proteins (Schiffer *et al.*, 1992).

The physical basis of the interaction of aromatic residues at the interfacial zones are believed to arise from a variety of amphipathic interactions related to the nature of the side groups, particularly the imino group of tryptophan. Suggestions range from hydrogen bonding effects (Schiffer *et al.*, 1992) to dipole and van der Waals interactions (Wimley and White, 1993). In a comprehensive study of four carefully chosen tryptophan analogues interacting with 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine membranes, it was found that the interactions were more general in nature involving aromaticity and the molecular shape of the amino acids (Yau *et al.*, 1998). It was suggested that the flat rigid shape of tryptophan limits access to the hydrocarbon core for entropic reasons while the electrical characteristics peculiar to aromatic molecules (their 'aromaticity'; Dougherty, 1996) make tryptophan ideally suited to residing in the electrostatically complex interface environment (Yau *et al.*, 1998). The effect of tryptophan on the positioning of transmembrane helices in biological membranes has been supported by studies using synthetic peptides in model systems (Braun and von Heijne, 1999). These studies have increased our understanding of membrane protein structure and have improved our ability to predict the tertiary structure of such proteins from their primary amino acid sequences.

1.2.1.2 Covalent Binding of Proteins to Membrane Lipids - Lipoproteins

Lipoproteins are proteins that are covalently linked to the lipid membrane by an N-terminal cysteine residue. They were first identified in Gram-negative bacteria, the classic example being the murein lipoprotein (MLP/LLP) of *E. coli* which was one of the earliest and best characterized (Braun and Hantke, 1974). Lipoproteins are synthesized as precursors with an N-terminal signal

peptide which directs the polypeptide to the secretory pathway prior to lipidation (Inouye *et al.*, 1983). At the cytoplasmic membrane, the cysteine residue of the lipoprotein consensus motif LXXC, is modified by the addition of diglyceride (Hantke and Braun, 1973), a process catalyzed by the membrane-embedded glyceryl transferase and *O*-acyl transferase (Tian *et al.*, 1989; Gan *et al.*, 1993; Gupta *et al.*, 1993). This is followed by cleavage of the signal peptide at the N-terminal side of the modified cysteine by a lipoprotein specific signal peptidase II (Innis *et al.*, 1984). The liberated amino group of cysteine is then further acylated by an *N*-acyl transferase (Lin *et al.*, 1978).

The targeting of a protein to the inner membrane of a Gram-negative bacterium is determined by the presence of an aspartic acid residue directly following the modified cysteine in the LXXC motif, while the presence of any other amino acid localizes the protein to the outer membrane (Yamaguchi *et al.*, 1988). In the case of MLP/LLP, the protein is trimerized in the periplasmic space before assembly in the outer membrane (Braun and Bosch, 1973; Braun *et al.*, 1976). It has been found that approximately one third of all lipoprotein is also covalently linked to the peptidoglycan (Braun and Rehn, 1969). The site of this linkage has been characterized in MLP as a C-terminal lysine which is amide linked to the carboxyl group of *m*-diaminopimelic acid within the *E. coli* cell wall peptides (see Section 1.2.2.1 below; Braun and Sieglin, 1970; Braun and Wolff, 1970; Braun, 1975).

A number of Gram-negative bacteria produce lipoproteins which have strong noncovalent interactions with peptidoglycan (Braun and Wu, 1994). These lipoproteins may mediate the interaction between the outer membrane and peptidoglycan, playing an important structural role. One such protein is the peptidoglycan-associated lipoprotein (PAL) of *E. coli* (Mizuno, 1979; Mizuno, 1981). PAL belongs to the Tol-PAL multiprotein complex which is involved in maintaining outer membrane integrity (Lazzaroni and Portalier, 1981; Lazzaroni and Portalier, 1992). Mutation of the *tol-pal* genes results in the formation of outer membrane vesicles, indicative of a defect in cell envelope assembly (Bernadac *et al.*, 1998). This system is composed of the TolA, TolQ and TolR protein which are located on the cytoplasmic membrane (Dérrouiche *et al.*, 1995; Journet *et al.*, 1999) and the periplasmic TolB which interacts with the outer membrane PAL (Bouveret *et al.*, 1995; Journet *et al.*, 1999). Not only do PAL and TolB interact with each other, PAL interacts with OmpA while TolB has been shown to interact with both OmpA and the major lipoprotein (Lpp) (Clavel *et al.*, 1998). It has thus been proposed that PAL and TolB proteins may be part of a larger complex involved with anchoring the outer membrane to the peptidoglycan (Bouveret *et al.*, 1999). Peptide competition studies have demonstrated that the region spanning from residue 89-130 of PAL, which contains a highly conserved α -helical domain (Koebnik, 1995), is sufficient to bind purified peptidoglycan or TolB. This interaction was found

to be exclusive, suggesting that PAL may exist in two forms in the cell envelope where one is bound to TolB and the other to peptidoglycan (Bouveret *et al.*, 1999). Surprisingly, while the structural importance of these lipoproteins in the Gram-negative cell structure is becoming more apparent, a gene (*slp*) of *B. subtilis* has been described which encodes for a putative lipoprotein which shows significant homology to the peptidoglycan-associated lipoproteins of *E. coli* and *Haemophilus influenzae* (Hemilä, 1991). However, interaction with the putative product of the *B. subtilis slp* gene and peptidoglycan, and any structural role the protein may play in Gram-positive cell envelope cohesion remains to be determined.

While lipoproteins from Gram-negative organisms have been extensively studied, examples from Gram-positive species have only relatively recently been identified. Identification of Gram-positive lipoproteins relied mainly on predictions from sequence analysis and the matching of the lipoprotein signal peptide cleavage site, L(S,A)(A, G) ↓ C(S,G), which seems to be conserved between Gram-positive and Gram-negative bacteria (von Heijne, 1989a; Braun and Wu, 1994; Sutcliffe and Russell, 1995). Other methods of identifying lipoproteins include radiolabelling the protein with [³H]-palmitic acid (Hayashi and Wu, 1990) or accumulation of lipoprotein precursors by the addition of globomycin, an inhibitor which specifically prevents lipoprotein processing by signal peptidase II (Hussain *et al.*, 1980; Dev, *et al.*, 1985).

It has been noted that proteins with related function which are to be found in the periplasm of Gram-negative bacteria are often lipid modified in Gram-positive bacteria (Gilson *et al.*, 1988). It has been suggested that the lipid moiety is required to anchor these proteins to the membrane surface in order to prevent their loss due to the porous nature of the Gram-positive peptidoglycan wall and lack of an outer membrane (Sutcliffe and Russell, 1995). In Gram-negative bacteria these soluble proteins are generally free in the periplasmic region, being physically retained by the inner and outer cell membranes. One such example is the β -lactamase (Bla) which confers resistance to the β -lactam antibiotics. Soluble and periplasmic in Gram-negative bacteria (Koshland and Botstein, 1980), the β -lactamases of *Bacillus licheniformis* and *Staphylococcus aureus* are lipoproteins (Nielsen *et al.*, 1981; Lai *et al.*, 1981; Nielsen and Lampen, 1982). While the membrane-associated Bla lipoprotein in *Bacillus* appears to be part of the secretion pathway for the production of extracellular β -lactamase (Nielsen and Lampen, 1982; Nielsen and Lampen, 1983), in some strains of *S. aureus* almost all of the enzyme is cell-associated (East and Dyke, 1989). In fact, a mutant Bla of *S. aureus* which is processed solely by type I signal peptidase and secreted, fails to protect the cells from β -lactam antibiotics (Navarre *et al.*, 1996). More recently the first lipidated β -lactamase (*bro*) has been identified on the outer membrane of the Gram-negative organism *Moraxella catarrhalis* (Bootsma *et al.*, 1999). Such Bro β -lactamase producing

strains of *M. catarrhalis* have only emerged in the past twenty years. A 31% G+C content for *bro* compared with the rest of the genome where the G+C content is 41%, strongly indicates that the Bro β -lactamase of *M. catarrhalis* is a recent acquisition from a Gram-positive source (Bootsma *et al.*, 1999).

Besides the role in conferring antibiotic resistance, lipoproteins of Gram-positive bacteria serve numerous functions ranging from substrate binding for nutrient uptake to adhesins and protein transport. Adhesion is an essential primary element for bacterial colonization and is of particular relevance in the oral cavity where there are numerous organisms which coaggregate and also interact with the available oral surfaces, namely the tooth and epithelial tissues. The FimA protein associated with the fimbriae of *Streptococcus parasanguis* FW213 mediates adhesion with saliva-coated hydroxyapatite (SHA) (Olingo and Fives-Taylor, 1993) as well as fibrin monolayers associated with endocarditis (Burnette-Curley *et al.*, 1995). It is a member of the streptococcal lipoprotein receptor antigen (Lra) group 1 which contains lipoproteins of approximate M_r 35 000 (Jenkinson, 1994). FimA shows 92% identity with PsaA of *Streptococcus pneumoniae* (Sampson *et al.*, 1994), 87% identity with the saliva binding protein (SsaB) of *S. sanguis* (Ganeshkumar *et al.*, 1991) and 86% identity with ScaA of *S. gordonii* PK488 (Kolenbrander *et al.*, 1994). The coaggregation adhesin, ScaA of *S. gordonii* is a 34 800 lipoprotein which mediates coaggregation with *A. naeslundii* PK606 (Kolenbrander and Andersen, 1990). Mutants of *S. gordonii* PK488 which do not coaggregate with *A. naeslundii* are also ScaA negative (Kolenbrander and Andersen, 1990). The importance of ScaA is evident from southern blot analysis where reactive chromosomal fragments are found in 19 streptococci with two *ScaA* probes including some non-coaggregating streptococci such as *S. salivarius* ATCC25975 (Kolenbrander *et al.*, 1994). The presence of *ScaA* homologues in non-coaggregating strains of streptococci is not unusual as several adhesins have been reported to serve several functions (Jenkinson *et al.*, 1993; Pancholi and Fischetti, 1992). Besides its role in adhesion, ScaA has also been reported to be part of the ATP binding-protein cassette (ABC) transporter system where it may function as a binding lipoprotein which recognizes small soluble molecules for transport (Kolenbrander *et al.*, 1994).

The ABC family of transporters found in bacteria (Higgins, 1992) is responsible for the transmembrane uptake of a wide range of substrates including saccharides, amino acids, peptides, vitamins, anions and cations (Tam and Saier, 1993). This system was first identified in Gram-negative bacteria as the binding-protein-dependent transport system (Higgins *et al.*, 1988; Hyde *et al.*, 1990; Nikaido and Saier, 1992) but later found to include the lipoprotein transporters of Gram-positive bacteria (Alloing *et al.*, 1990; Russell *et al.*, 1992; Alloing, *et al.*, 1994). The ABC transport system is comprised of three components: an ATP-binding protein (ATPase), a hydrophobic membrane protein and a substrate-specific binding protein. The genes coding for

ABC transport proteins are generally organized in an operon (Furlong, 1987; Alloing *et al.*, 1990). In Gram-negative bacteria the binding proteins are usually found free in the periplasm, but in Gram-positive bacteria they are membrane anchored lipoproteins (Gilson *et al.*, 1988; Alloing *et al.*, 1994). Several Gram-positive transport systems and multi-purpose lipoproteins have now been identified. These include the AmiA, AliA and AliB lipoproteins of the oligopeptide permease system of *S. pneumoniae* which although located on two separate operons, are highly homologous (Alloing *et al.*, 1990; Alloing *et al.*, 1994). Each binding protein appears to have a slightly different specificity for oligopeptides ranging from two to seven residues in length. In combination, these binding proteins allow a wide range of substrates to be transported by the Ami system (Alloing *et al.*, 1994). In addition to their binding protein function, mutations in *amiA* and *aliA* cause a 50% reduction in pneumococcal cell binding to epithelial and endothelial tissues (Cundell *et al.*, 1995).

The three HppA, G and H peptide-binding lipoproteins of *S. gordonii* are genetically and structurally similar to the Ami-like permeases of *S. pneumoniae* (Jenkinson, 1992; Jenkinson *et al.*, 1996). However unlike the Ami system where the three associated peptide-binding proteins have overlapping substrate specificities (Alloing *et al.*, 1994), both HppA and HppH proteins are essential for oligopeptide uptake, primarily hexa- and heptapeptides. Thus the HppA and HppH proteins may interact with each other or the integral membrane components of an ABC transport system necessary for oligopeptide uptake (Jenkinson *et al.*, 1996). In addition, *hppA* mutants of *S. gordonii* are also retarded in their ability to adhere to strains of *A. naeslundii* and to aggregate in human saliva or serum (Jenkinson and Easingwood, 1990). It is becoming generally apparent that these binding-protein dependent permeases not only act as transporters, but also play a role either directly or indirectly in adhesion.

1.2.2 Attachment of Proteins to Peptidoglycan

1.2.2.1 Synthesis and Structure of Peptidoglycan

Gram-positive bacteria are morphologically simple cells comprised of a cytosol, a single cytoplasmic membrane and a cell wall (Giesbrecht *et al.*, 1976). The cell wall not only provides a rigid structure for protection against mechanical and osmotic stress (Salton, 1952; Salton, 1994), but also acts as a site for protein attachment. The cell wall of Gram-positive bacteria is composed of peptidoglycan with attached accessory molecules such as teichoic acid, polyphosphates and carbohydrates (Salton, 1994; Hancock, 1997). The peptidoglycan heteropolymer is found exclusively in prokaryotes and is composed of substituted sugar units in chains. These glycan strands are linked by peptide cross bridges. The glycan strands of the cell wall consist of the repeating disaccharide *N*-acetylmuramic acid-(β 1-4)-*N*-acetyl-

glucosamine (MurNAc-GlcNAc) (Ghuysen, 1963a, b) of approximately 5 to 30 subunits depending on the bacterial species (Glauner *et al.*, 1988; Snowden and Perkins, 1990). Usually the D-lactyl moiety of MurNAc is amide linked to the short peptide component of peptidoglycan (Munoz *et al.*, 1966; Tipper *et al.*, 1967) which is then cross-linked to other peptides attached to the neighbouring glycan strand (Tipper and Strominger, 1968; Tipper and Berman, 1967) providing the three-dimensional network of the murein exoskeleton (Labischinski and Maidhoff, 1994).

Cell wall synthesis occurs in the cytoplasm, the plasma membrane and the cell wall itself (Strominger, 1968). Synthesis of the building blocks of peptidoglycan begins in the cytoplasm with the attachment of N-acetylglucosamine (GlcNAc) to uridine diphosphate (UDP) to form UDP-N-acetylglucosamine (UDP-GlcNAc). UDP-GlcNAc and phosphoenolpyruvate (PEP) are then converted to UDP-N-muramic acid (UDP-MurNAc) (Strominger, 1959; Gunetileke and Anwar, 1966). Five amino acids are joined sequentially to UDP-MurNAc to form the pentapeptide (Ito and Strominger, 1964; Nathensen *et al.*, 1964; Mizuno *et al.*, 1973). The synthesis of each amide bond within the wall peptide consumes 1 high-energy phosphate (ATP) resulting in UDP-MurNAc-L-Ala-D-isoGlu-L-Lys-D-Ala-D-Ala (UDP-MurNAc-pentapeptide) (van Heijenoort, 1994; Park, 1952; Chatterjee and Park, 1964).

The synthesis of wall peptides occurs at the membrane where the UDP-MurNAc-pentapeptide is linked to a C_{55} -pyrophosphate carrier molecule resulting in the formation of C_{55} -PP-MurNAc-pentapeptide or lipid I (Anderson and Strominger, 1965). The disaccharide unit is formed next with the addition of UDP-GlcNAc to the muramoyl moiety generating the disaccharide lipid II precursor, C_{55} -PP-MurNAc-(pentapeptide)- β -1-4-GlcNAc (Higashi *et al.*, 1967; Higashi *et al.*, 1970). Lipid II is further modified with the addition of amino acid cross-bridge units to the pentapeptide moiety by specialized tRNA species (Anderson *et al.*, 1967; Stewart *et al.*, 1971; Green and Vold, 1993). In *Staphylococcus* five glycine residues are linked to the L-lysine wall peptide (Roberts, 1974; Roberts *et al.*, 1974). The final lipid II precursor, C_{55} -PP-MurNAc-L-Lys-D-iso-Glu-l-Lys(Gly₅)-D-Ala-D-Ala)- β 1-4-GlcNAc, is then translocated across the cytoplasmic membrane where it is incorporated into the peptidoglycan. This assembly is catalyzed by Class A penicillin binding proteins (PBPs) (Ghuysen, 1991; Goffin and Ghuysen, 1998) which are responsible for the polymerization of glycan from the disaccharide precursor as well as the transpeptidation (cross-linking) of wall-peptides (Nakagawa *et al.*, 1984). Transpeptidation results from the proteolytic removal of the D-alanine at the C-terminal end of the pentapeptide and the formation of a new amide bond between the amino group of the cross-bridge and the carbonyl of D-alanine (Tipper *et al.*, 1967).

The amino acid composition of peptide cross-bridges can differ between bacterial species, giving rise to a variety of peptidoglycans (Fig 1.3). In Gram-negative bacteria such as *E. coli* and the Gram-positive *Listeria monocytogenes*, the cross-link consists of peptide bonds between the carbonyl group of D-alanine and the amino group of *m*-diaminopimelic acid which is a precursor of lysine (Schleifer and Kandler, 1972). In most Gram-positive organisms L-lysine replaces *m*-diaminopimelic acid. The L-lysine of the wall peptides are then amide linked to carbonyl of D-alanine in the wall peptide of a neighbouring strand. In *S. aureus* the wall peptides are joined by 5 glycine residues (pentaglycine), and in *Streptococcus pyogenes* by 2 alanines (Fig 1.3; Schleifer and Kandler, 1972).

1.2.2.2 Covalent Binding of Proteins to Gram-positive Cell Walls

A vast number of surface proteins of Gram-positive bacteria have now been cloned and sequenced (Fischetti *et al.*, 1990). These surface proteins mediate the interaction between the bacterium and its external environment such as those following contact with host tissues (Jenkinson, 1995). The functions of surface proteins are not only varied but often vital for bacterial survival and often contribute to bacterial virulence. For example, surface proteins are involved in nutrient binding (Sato *et al.*, 1997), nutrient metabolism (Burne and Penders, 1992; Colby *et al.*, 1995; Ohnishi *et al.*, 1995), bacterial aggregation (Douglas, 1995; Dunny and Leonard, 1997), adhesion (Whittaker *et al.*, 1996) and evasion of the host immune system (Lee, 1995).

The earliest attempts at investigating attachment of proteins to Gram-positive cell walls employed the use acid extraction, proteases, detergents and boiling. However, these methods resulted in extremely low protein yields as well as protein hydrolysis (Lancefield, 1962). As a result, muralytic enzymes which hydrolyze the cell wall peptidoglycan were used effectively to solublize surface protein intact. For example, lysostaphin, a bacteriolytic enzyme secreted by *Staphylococcus simulans* (Schindler and Schuhardt, 1964) that digests the pentaglycine cross-bridge, and lysozyme, an N-acetyl-muramidase, were initially used to release protein A from *S. aureus* (Sjöquist *et al.*, 1972a, b). When the two extracts were subject to SDS-PAGE, it was found that the lysozyme-released protein A migrated more slowly (Sjöquist, 1972b). Acid hydrolysis of the lysozyme-released protein A identified the presence of GlcNAc and MurNAc, indicating that protein A must be linked to the *S. aureus* cell wall (Sjöquist, 1972b).

In later studies, the cell-associated region of the streptococcal M6 protein was isolated by trypsin digestion (Pancholi and Fischetti, 1988). The protected wall-associated domain was then released by phage lysin digestion of the peptidoglycan. Antibody probes were used to identify and purify the 16 000 C-terminal fragment. Amino acid sequence analysis

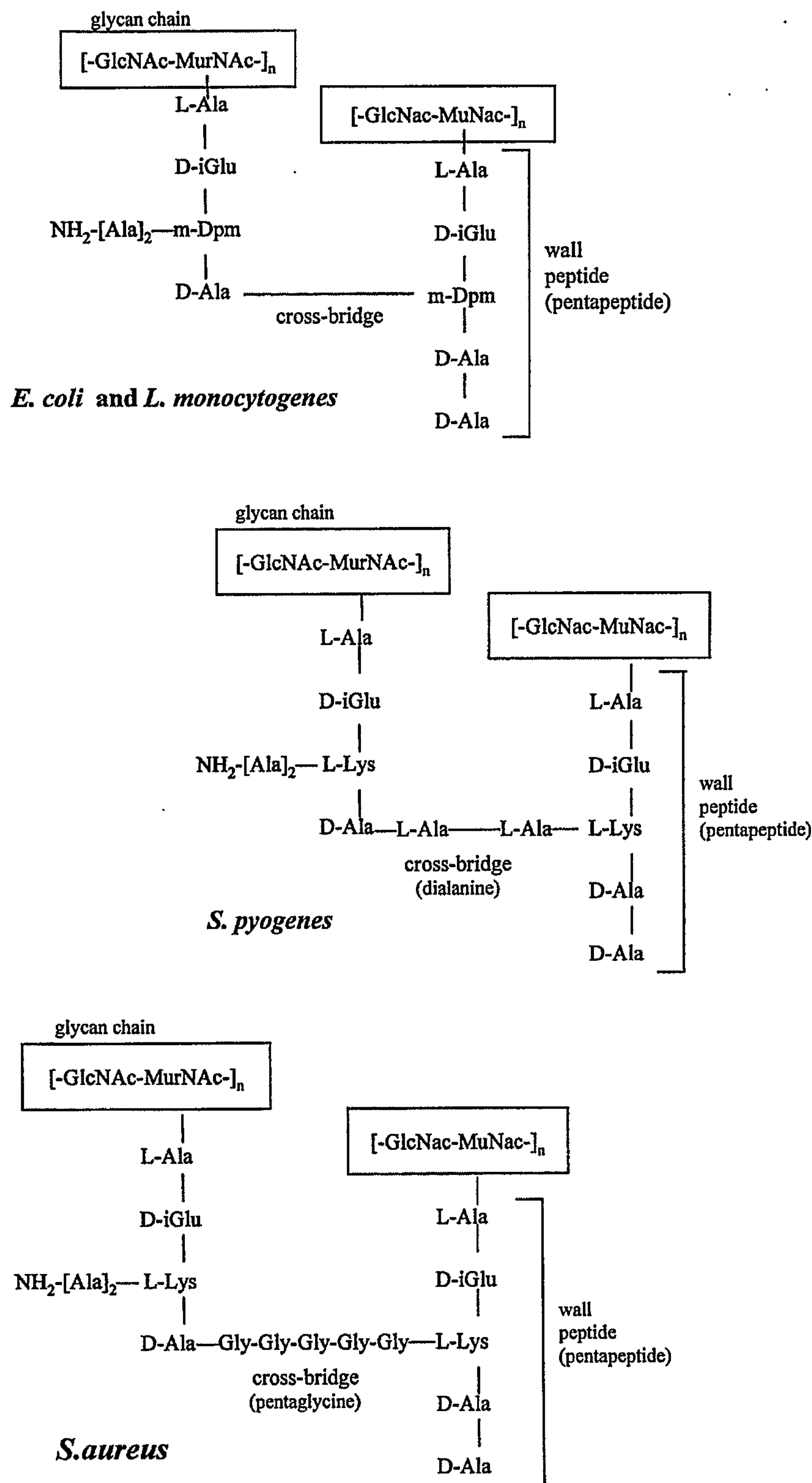


Figure 1.3 Diagram of the peptidoglycan structures of *E. coli*, *L. monocytogenes*, *S. pyogenes* and *S. aureus*. The glycan chains are composed of repeating disaccharides, GlcNAc-MurNAc. The wall pentapeptide is amide linked to MurNAc. Adjacent wall peptides are cross-linked by a variety of amino acid cross-bridges; direct D-alanine-*m*-diaminopimelic acid (m-Dpm) in *E. coli* and *L. monocytogenes*, by dialanine in *S. pyogenes* and by pentaglycine in *S. aureus*.

revealed that the isolated C - terminal peptide contained residues from amino acid 289 within the M6 sequence and that the hydrophobic domain was absent.

Covalent linkage or sorting of surface proteins to the Gram-positive cell wall peptidoglycan was subsequently investigated with staphylococcal protein A (Schneewind *et al.*, 1993). Lysostaphin which cleaves the pentaglycine cross-bridge of the staphylococcal peptidoglycan, solubilized protein A as a uniform species on SDS-PAGE (Schneewind *et al.*, 1992). In contrast, a muramidase that cleaved the glycan strands released protein A as a spectrum of bands which migrated slower than the lysostaphin-released protein A (Schneewind *et al.*, 1993). The difference in mobility was suspected to be a result of the differences in linked peptidoglycan moieties. This was confirmed when muramidase-released protein A was digested with lysostaphin, resulting in a uniform product with the same mobility as the lysostaphin-released protein (Schneewind *et al.*, 1993).

Subsequent sequence comparisons and analysis of surface-attached proteins have resulted in the identification of a common C-terminal arrangement of amino acids (Fischetti *et al.*, 1991). Terminating with a small charged tail, a 15-30 amino acid hydrophobic region is preceded by a highly conserved pentapeptide, the LPXTG motif, which lies next to a 50-125 amino acid wall-spanning domain (Fischetti *et al.*, 1990; Schneewind *et al.*, 1991). The wall spanning region is characterized by the high percentage of proline-glycine (10-32%) and threonine-serine (19-38%) (Fig 1.4; Pancholi and Fischetti, 1988; Fischetti *et al.*, 1991).

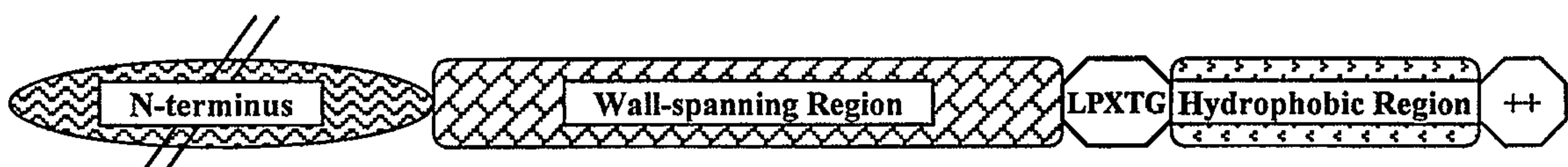


Figure 1.4 C-terminal domains in covalently bound cell surface proteins of Gram-positive bacteria.

The role of the LPXTG motif, hydrophobic domain and charged tail of the staphylococcal protein A sorting signal was examined using deletion mutations (Schneewind *et al.*, 1992). C-terminal deletions of the charged tail and hydrophobic domain as well as the

entire sorting signal, resulted in secretion of streptococcal protein A. Additionally, when the C-terminal sorting signal of the staphylococcal protein A is fused to the normally secreted staphylococcal enterotoxin B (Seb), β -lactamase (BlaZ), or *E. coli* alkaline phosphatase (PhoA), the hybrid proteins become anchored to the cell wall (Schneewind *et al.*, 1992; Schneewind *et al.*, 1993). It is apparent that the entire sorting signal is required for surface attachment. More recent emerging evidence indicates that the C-terminal hydrophobic domain and the charged tail act to restrain surface proteins during translocation across the cytoplasmic membrane, providing the opportunity for the proteolytic cleavage and covalent attachment of the protein to the cell wall peptidoglycan (Navarre and Schneewind, 1999).

Direct evidence for the involvement of the LPXTG motif in anchoring of surface proteins to Gram-positive cell wall peptidoglycan came from studies using hybrid proteins (Navarre and Schneewind, 1994). These hybrid proteins contained the *S. aureus* protein A sorting signal flanked by indicator sequences. These proteins were organized such that the N-terminus consisted of an N-terminal portion of the enterotoxin B peptide (Seb) attached to the sorting signal composed of the LPXTG motif, an hydrophobic domain and a charged tail, followed by the mature form of either the *E. coli* maltose-binding protein (MalE) or the staphylococcal β -lactamase, BlaZ.

When expressed in *S. aureus*, these hybrid proteins were exported from the cytoplasm and cleaved into two products. The cellular location of the two fragments were determined by cell fractionation and protease protection assays. It was found that the N-terminal Seb fragment was located in the cell wall compartment. This fragment was sensitive to extracellular proteases, indicating that the peptide was successfully translocated. The fractionation experiments determined that the C-terminal fragments (BlaZ or MalE) were predominantly in the cytoplasmic or membrane compartments. The site of proteolytic cleavage was determined by sequencing the MalE fragment. The N-terminal residue was found to be the glycine of the LPXTG motif. It was thus speculated that a covalent linkage most likely occurred between the carbonyl-group of the liberated threonine and the amino-group of the peptide cross-bridge glycine of the peptidoglycan (Navarre and Schneewind, 1994). This was confirmed in a subsequent study wherein a hybrid protein composed of the Seb signal sequence fused to the mature MalE and attached to a trypsin cleavage site located next to the sorting signal of protein A was expressed in *S. aureus* (Schneewind *et al.*, 1995). The lysostaphin-released hybrid was purified by high affinity chromatography and reverse-phase high-performance liquid chromatography. Trypsin treated and untreated fragments were subject to electrospray mass spectrometry and the exact peptide masses determined. The mass differences were consistent with the presence of three glycines linked to the threonyl of

LPXTG. These measurements demonstrate that the hybrid protein is linked to the staphylococcal cell wall through an amide bond between the C-terminal carbonyl of threonine and the amino acids of the pentaglycine cross-bridge (Fig 1.5). While Gram-positive bacteria display variation in the composition of their peptidoglycan cross-bridges (Fig1.3), the presence of a free amino group is a common feature (Schleifer and Kandler, 1972). It is therefore expected that sorting of surface proteins harbouring the LPXTG motif occurs by a universal mechanism (Schneewind *et al.*, 1995).

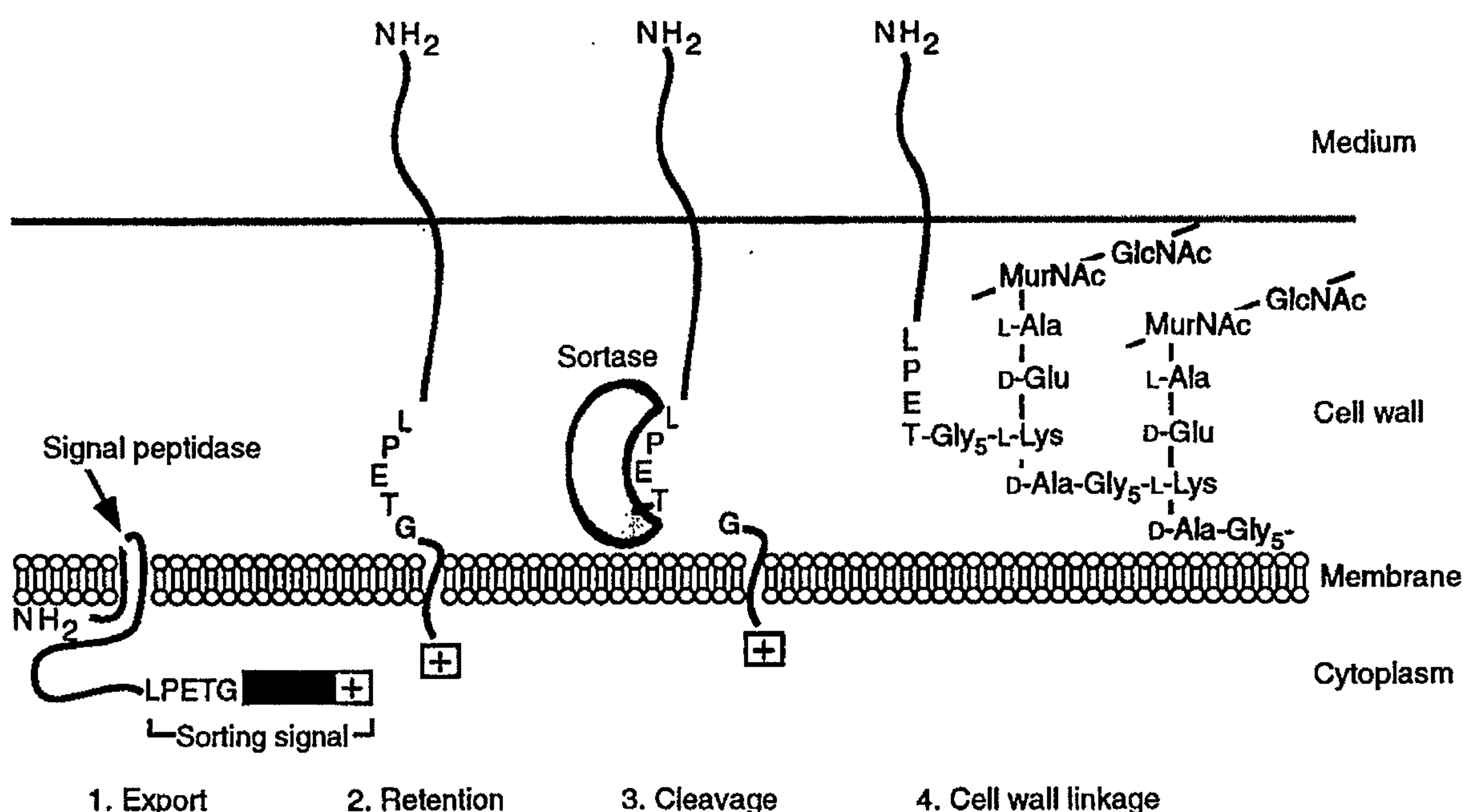


Figure 1.5 Four distinct stages in the cell-wall anchoring of surface proteins in *S. aureus* (after Schneewind *et al.*, 1995)

1.2.3 The Transpeptidase (Sortase) Responsible for Processing the LPXTG Site

Evidence for the covalent attachment of surface proteins to the cell wall of Gram-positive bacteria via the LPXTG motif instigated the search for the transpeptidase responsible for cleavage between the threonine and glycine residues and the subsequent attachment of the surface proteins to the amino acid cross bridge of peptidoglycan. The transpeptidase, or sortase, responsible for the anchoring of surface proteins in *S. aureus*, was identified in

temperature-sensitive mutants which had been transformed with a hybrid reporter protein composed of the enterotoxin B peptide and the sorting signal peptide from *S. aureus* protein A (Seb-Spa₄₉₀₋₅₂₄; 1.2.2.2; Schneewind *et al.*, 1993). The processing and anchoring of the Seb-Spa₄₉₀₋₅₂₄ reporter protein to the cell wall of *S. aureus* normally takes 2 min, but was significantly delayed in two *S. aureus* mutants (Mazmanian *et al.*, 1999). Since the delay in protein sorting might have been caused by a variety of defects other than one in the sortase itself, care was taken to establish that the mutants were not altered in cell wall cross-bridge formation, cell wall synthesis or in the actual anchor structure of the surface proteins itself. The surface protein sorting A gene (*srtA*) was subsequently isolated from a *S. aureus* chromosomal plasmid library capable of complementing the sorting defect in the temperature-sensitive mutants (Mazmanian *et al.*, 1999). The *srtA* gene was found to code for a 206 amino acid protein with a N-terminal membrane anchor. Database searches found *srtA* homologues in a wide variety of Gram-positive bacteria including *A. naeslundii*, *B. subtilis* and *S. mutans* (Mazmanian *et al.*, 1999).

Analysis of the deduced amino acid sequence of the *srtA* gene and its homologues showed the presence of a conserved cysteine at position 184 (Mazmanian *et al.*, 1999). While the presence of the conserved cysteine implied a role at the active site of the sortase, transpeptidation is also known to occur via active site hydroxyl groups (Rasmussen and Strominger, 1978). To distinguish between these two possibilities, inhibition of the anchoring mechanism was determined in the presence of specific inhibitors. The study found that cleavage of the LPXTG motif is sensitive to methanethiosulphates and *p*-hydroxymercuribenzoic acid indicating that the sortase must be a sulfidryl-containing enzyme (Ton-That and Schneewind, 1999). The importance of C¹⁸⁴ was definitively shown by construction of the mutated enzyme C184A which was found to be inactive (Ton-That and Schneewind, 1999). In addition, the formation of an acyl-enzyme intermediate was identified with the use of hydroxylamine, a strong nucleophile which attacks thioester bonds (Lawrence and Strominger, 1970; Ton-That *et al.*, 1999). Electrospray ionization mass spectrometry of purified surface protein peptides released in the presence of hydroxylamine showed the presence of hydroxamate at the threonine of the LPXTG motif. These data suggest that SrtA cleaves LPXT↓G to form a transient covalent intermediate between the threonine and the sulfidryl group of a cysteine residue in the sortase (Surface Protein-LPXT(CO-S)-SrtA), prior to this intermediate being resolved by the nucleophilic attack of an amino group within the pentaglycine cross-bridge within the peptidoglycan. Evidence from antibiotic inhibition studies indicate that the sortase utilizes the peptidoglycan precursor, lipid II (1.2.2.1), as the substrate for binding rather than the mature assembled cell wall (Ton-That and Schneewind,

1999). Incorporation of this surface protein-lipid II intermediate would result in the final attachment of the protein to the cell wall (Fig 1.5; Ton-That *et al.*, 1999).

1.2.4 An Alternative Cell Surface Protein Binding Mechanism in Gram-positive Bacteria

Until recently, the majority of surface proteins of Gram-positive bacteria were believed to be anchored to the cell wall peptidoglycan by the C-terminal LPXTG consensus motif (1.2.2.2) or to the cytoplasmic membrane by an N-terminal lipoprotein anchor identifiable by the LXXC consensus motif (1.2.1.2). However, a novel mechanism of cell attachment mediated by electrostatic interactions has been reported for the BspA surface protein of *Lactobacillus fermentum* BR11 which lacks both of these processing sites (Turner *et al.*, 1997).

Sequence analysis showed that BspA possessed 88.6% identity with Cnb, a collagen binding protein of *Lactobacillus reuteri* which belongs to the family III solute binding proteins of ABC-type transporter systems (1.2.1.2). Alignment of BspA with *E. coli* proteins showed that it was similar to the L-cystine binding protein, FliY. Insertional inactivation of *bspA* in *L. fermentum* resulted in impaired L-cystine uptake and defective growth in the presence of oxygen as well as sensitivity to paraquat (Turner *et al.*, 1999).

As described above (1.2.1.2), ABC-binding proteins of Gram-positive bacteria are usually lipoproteins. However, the BspA of *L. fermentum* lacks the signal peptidase II consensus LXXC motif, characteristic of lipoproteins. Analysis of BspA showed that the deduced amino acid sequence possessed a highly alkaline pI of 10.59, suggesting that the protein may be anchored electrostatically to negatively charged components of cell wall envelope such as teichoic acid (Turner *et al.*, 1997). Evidence in support of this model came from the observations that treatment of cells with 5M LiCl or glycine hydrochloride (at pH < 2.0) selectively removed BspA from the cell surface. The result not only indicated that the protein was surface located, but that electrostatic rather than covalent forces were involved in attachment of this protein to the cell envelope of *L. fermentum* (Turner *et al.*, 1997).

1.2.5 Aims of the Present Study

The main focus of the research was to determine the precise mode of surface attachment employed by the Ftf of *S. salivarius*. Sequence analysis of the *ftf* gene should identify if the Ftf were attached to the membrane (membrane anchored or lipoprotein), covalently linked to the peptidoglycan, localised by ionic interactions to secondary wall polymers or by a yet unidentified and novel mechanism.

CHAPTER 2

ROUTINE METHODS

This chapter describes methods used routinely during this research. Experimental methods that were developed during the course of this study are described in the relevant chapters.

2.1 REAGENTS, CHEMICALS AND ENZYMES

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

Ampicillin, tetracycline, cesium chloride, deoxyribonuclease (DNase) I, lysozyme, ribonuclease (RNase) A, exonuclease III, S1 nuclease, Klenow DNA polymerase I, 5-bromo-4-chloro-3-indolyl-galactoside (X-gal): Boehringer Mannheim, Sydney, NSW, Australia.

Tryptone, yeast extract, Columbia blood agar base, Todd Hewitt broth, Brain Heart Infusion broth, bacteriological agar, hydrolyzed casein: Oxoid, Basingstoke, UK.

[U-¹⁴C-fructosyl]-labelled sucrose, [³⁵S]-dATP: Dupont NEN, Boston, Mass, USA.

xylene cyanol FF: International Biotechnologies Inc., Conn, USA.

Ultima GoldTM scintillation fluid: Packard Instrument Co., Il, USA

Chloramphenicol, polyethyleneglycol, iodoacetamide, guanine, riboflavin, biotin, folic acid, *p*-aminobenzoic acid, acquacide III: Calbiochem-Novabiochem, Melbourne, VIC, Australia.

Defibrinated horse blood: AMYL Media Pty Ltd, Sydney, NSW, Australia.

Horse serum (inactivated): Commonwealth Serum Laboratories, Melbourne, VIC, Australia.

Whatman 3M filter paper, Whatman GAF/B glass fibre filters: Whatman, International Ltd., Madison, UK.

Millipore Type VS filter: Millipore, North Ryde, NSW, Australia.

Miniskirt filter: Sartorial Australia Pty. Ltd., Oakleigh, VIC, Australia.

Polyvinylidene difluoride membranes: BioRad Laboratories Pty. Ltd., Regents Park, NSW, Australia.

Light paraffin oil, isopropanol: Merck, Melbourne, VIC, Australia.

Electrophoresis gel stain solution ('Gradipure' stain): Gradipure Ltd, Sydney, NSW, Australia.

Coomassie® Plus Protein Assay Reagent: Pierce, IL, USA.

Wizard™ Plus Miniprep DNA Purification Kit, the Transformer™ Site-Directed Mutagenesis Kit : Promega, WI, USA.

Agarose NA, bind-silane, repel-silane, deoxynucleotide phosphates (dNTPs), Vent DNA polymerase, T7 DNA polymerase, T7 Sequencing kit: Pharmacia Biotech, Melbourne, VIC, Australia.

Acrylamide, adenosine 5'-triphosphate (ATP) (sodium salt), ammonium persulfate, bromophenol blue, calcium chloride, dithiothreitol (DTT), ethidium bromide, ethylenediaminetetraacetic acid (EDTA), ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), formaldehyde, fructose, glucose, L-histidine, β-mercaptoethanol, N,N'-methylene-bis-acrylamide, periodic acid, sodium piperazine, phenylmethylsulfonyl fluoride (PMSF), erythromycin, chloroform, proteinase K, phenol, isoamylalcohol, uridine, triphenyltetrazolium chloride, potassium phosphate dibasic, potassium phosphate monobasic, raffinose, fushin sulfite (Schiff's reagent), sodium azide, sodium dodecyl sulfate (SDS), sodium fluoride, sucrose, N,N,N',N'-tetramethyl-ethylenediamine (TEMED), isopropyl β-D-thiogalactopyranoside (IPTG), thiomersal, thioglycolate, amido black, adenine, uracil, cysteine, thiamine, pyridozamine, calcium pantothenate, triton X-100, trizma base (Tris), aprotinin, chymostatin, leupeptin, trypsin inhibitor, nicotinamide, benzamide: Sigma, Sydney, NSW, Australia.

Restriction enzymes, λ DNA-*Hind*III digest or λ DNA-*Hind*III digest/φχ-174RF DNA-*Hinc*II digest molecular weight markers, T4 nucleotide kinase, T4 DNA polymerase, T4 DNA ligase, calf-intestinal alkaline phosphatase: either Boehringer Mannheim, Sydney, NSW, Australia; Promega Corp., Sydney, NSW, Australia; New England Biolabs Arundel, QLD, Australia; or Pharmacia Biotech, Melbourne, VIC, Australia.

Broad range SDS-PAGE standards, gel filtration chromatography standards: Bio-Rad, Sydney, NSW, Australia. Non-denaturing-PAGE standards: Sigma, Sydney, NSW, Australia.

Oligonucleotides were synthesized on a Gene Assembler Plus (Pharmacia Biotech, Melbourne, VIC, Australia), using Pharmacia chemicals or purchased from Auspep Ltd, Melbourne, VIC, Australia or Beckman Instruments, Sydney, NSW, Australia.

Mutanolysin was kindly supplied by Dr K Yokogawa, Dainippon Pharmaceutical Co., ltd., Osaka, Japan, or alternatively, purchased from Sigma, Sydney, NSW, Australia.

2.2 BACTERIAL STRAINS, BACTERIOPHAGE, PLASMIDS AND PHAGEMIDS

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

TABLE 2.1. Bacterial strains, bacteriophages, plasmids and phagemids

Bacteria, Phage, Plasmid or Phagemid	Description ^{a, b}	Host Organism(s)	Source/Reference
Bacterium: <i>Streptococcus salivarius</i> ATCC 25975	-----	---	American Type Culture Collection (Hamilton, 1968)
<i>Streptococcus gordonii</i> LGR2	-----	---	Wyatt <i>et al.</i> , 1988
<i>Escherichia coli</i> LE392		---	Murray <i>et al.</i> , 1977
<i>Escherichia coli</i> NM522	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> F' <i>lacI^q</i> Δ (<i>lacZ</i>) <i>M15 proA⁺B⁺/supE thi</i> Δ (<i>lac-proAB</i>) Δ (<i>hsdMS-mcrB</i>)5 (r _K ⁻ m _K ⁻ McrBC ⁻)	---	Gough and Murray, 1983
<i>Escherichia coli</i> JM109	F ⁻ <i>traD36 proA⁻B⁺ lacI^q Δ(lacZ)M15/ Δ(lac-proAB) ginV44 e14⁻ gyrA96 recA1 relA1 endA1 thi hsdR17</i>	---	Yanisch-Perron <i>et al.</i> , 1985
<i>Escherichia coli</i> CJ236	F ⁻ <i>cat</i> (<i>pCJ105 = pOX38::cat = FΔ(HindIII)::cat</i> [<i>Tra⁺ Pil⁺ Cam^R/ ung-1 relA1 dut-1 thi-1 spoT1 thi supE Δ(lac-proAB)</i> [<i>mutS::Tn10</i>] [<i>F' proAB lacI^q ZΔM15</i>]	---	Kunkel <i>et al.</i> , 1987
<i>Escherichia coli</i> BMH 71-18 <i>mut S</i>		---	Zell and Fritz, 1987
Bacteriophage: λ L47.1 λ A-39	(srI λ 1-2) ^v <i>imm434cI⁻ NIN5 chiA131</i> λ L47.1 with Ftf encoding 8.2 kbp <i>Sau3A1</i> partial fragment of <i>S. salivarius</i> ATCC 25975 chromosomal DNA	<i>E. coli</i> <i>E. coli</i>	Loenen and Brammer, 1980 Pitty <i>et al.</i> , 1989
M13K07	M13 derivative carrying the mutation Met40Ile in gII, the origin of replication from P15A and the <i>Km^R</i> gene from Tn903 both inserted in the M13 origin of replication	<i>E. coli</i>	IBI Corporation (Vieira and Messing, 1987)
Plasmid/Phagemid: pIBI30 pIBI31 pVA838 pKRK102 pKRK104 pKRK105 pKRK106	Ap ^r (E) Ap ^r (E) Em ^r (E,S); Cm ^r (E) Ap ^r (E); Ftf ⁺ ; 4.6 kbp <i>Bam</i> H1 of λ A-39 cloned into pIBI31 Ap ^r (E); Ftf ⁺ ; 3.6 kbp <i>Eco</i> R1/ <i>Bgl</i> II of λ A-39 cloned into pIBI31 Ap ^r (E); Ftf ⁺ ; 0.6 kbp <i>Bgl</i> II/ <i>Eco</i> R1 of λ A-39 cloned into pIBI31 Ap ^r (E); Ftf ⁺ ; 2.6 kbp <i>Bgl</i> II/ <i>Bam</i> H1 of λ A-39 cloned into pIBI31	<i>E. coli</i> <i>E. coli</i> <i>E. coli</i> , <i>S. gordonii</i> <i>E. coli</i> <i>E. coli</i> <i>E. coli</i> <i>E. coli</i>	IBI Corporation IBI Corporation. Clewell <i>et al.</i> , 1982 This study This study This study This study

TABLE 2.1 (cont)

Plasmid or Phagemid	Description ^{a, b}	Host Organism(s)	Source/Reference
pKRK107	Ap ^r (E); Ftf ⁺ ; 0.5 kbp <i>Bam</i> H1/ <i>Eco</i> R1 of λ A-39 cloned into pIBI31	<i>E. coli</i>	This study
pKRK108	Ap ^r (E); Ftf ⁺ ; 3.9 kbp <i>Eco</i> R1/ <i>Eco</i> R1 of λ A-39 cloned into pIBI31	<i>E. coli</i>	This study
pKRK1969	Ap ^r (E); Ftf ⁺ ; deletion of pKRK104 from <i>Xba</i> I so as to leave 3.29 kbp of insert	<i>E. coli</i>	This study
pKRK1914	Ap ^r (E); Ftf ⁺ ; deletion of pKRK104 from <i>Xba</i> I so as to leave 2.97 kbp of insert	<i>E. coli</i>	This study
pKRK1816	Ap ^r (E); Ftf ⁺ ; deletion of pKRK104 from <i>Xba</i> I so as to leave 2.67 kbp of insert	<i>E. coli</i>	This study
pKRK1661	Ap ^r (E); Ftf ⁺ ; deletion of 1.44 kbp <i>Hind</i> III fragment from pKRK104	<i>E. coli</i>	This study
pKRK2969	Em ^r (E,S); Ftf ⁺ ; 3.29 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1969 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study
pKRK2914	Em ^r (E,S); Ftf ⁺ ; 2.97 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1914 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study
pKRK2816	Em ^r (E,S); Ftf ⁺ ; 2.67 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1816 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study
pKRK3969	Ap ^r (E); Ftf ⁺ ; deletion of 0.8 kbp <i>Eco</i> R1/ <i>Eco</i> RV fragment from pKRK1969	<i>E. coli</i>	This study
pKRK1001	Ap ^r (E); site directed mutated pKRK1969 coding for S880STOP	<i>E. coli</i>	This study
pKRK1002	Ap ^r (E); site directed mutated pKRK1969 coding for K636P, S875R and S877P	<i>E. coli</i>	This study
pKRK1003	Ap ^r (E); site directed mutated pKRK1969 coding for G696P, S875R and S877P	<i>E. coli</i>	This study
pKRK1004	Ap ^r (E); site directed mutated pKRK1969 coding for K739P, S875R and S877P	<i>E. coli</i>	This study
pKRK1005	Ap ^r (E); pKRK1002 with <i>ffl</i> deleted of region coding for Asp-635 to Ser-875	<i>E. coli</i>	This study
pKRK1006	Ap ^r (E); pKRK1003 with <i>ffl</i> deleted of region coding for Asp-695 to Ser-875	<i>E. coli</i>	This study
pKRK1007	Ap ^r (E); pKRK1004 with <i>ffl</i> deleted of region coding for Asp-738 to Ser-875	<i>E. coli</i>	This study
pKRK2001	Em ^r (E,S); 3.29 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1001 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study
pKRK2002	Em ^r (E,S); 3.29 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1005 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study
pKRK2003	Em ^r (E,S); 3.29 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1006 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study
pKRK2004	Em ^r (E,S); 3.29 kbp <i>Eco</i> R1/ <i>Sph</i> I of pKRK1007 in 7.0 kbp <i>Eco</i> R1/ <i>Sph</i> I of pVA838	<i>E. coli</i> , <i>S. gordonii</i>	This study

^a Ap^r, ampicillin resistance; Em^r, erythromycin resistance; Cm^r, chloramphenicol resistance. The letters in parentheses indicate whether or not the resistance genes function in *E. coli* (E) or *S. gordonii* (S).
^b Plasmids and phagemids were expressed in *Escherichia coli* NM522 or *S. gordonii* LGR2 as described (2.3).

2.3 MEDIA AND GROWTH CONDITIONS

Escherichia coli strains (Table 2.1) were grown aerobically in liquid culture by vigorously shaking in an appropriately sized conical flask at 37°C. The medium used was either '2xYT' medium (10 g l⁻¹ yeast extract, 10 g l⁻¹ tryptone, 5 g l⁻¹ NaCl; Sambrook *et al.*, 1989) or Luria-Bertani (LB) medium (5g l⁻¹ yeast extract, 10 g l⁻¹ tryptone, 5 g l⁻¹ NaCl; Sambrook *et al.*, 1989) supplemented with 100 µg ampicillin ml⁻¹, 500 µg erythromycin ml⁻¹ or 20 µg chloramphenicol ml⁻¹ as appropriate. Selection of *E. coli* strains harbouring genes cloned into pIBI vectors (Table 2.1) was accomplished on agar plates made from one or the other medium and containing 1 mM isopropyl β-D-thiogalactoside (IPTG) and 100 µg 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X- gal) ml⁻¹ to allow for blue/white colony screening.

Liquid lysates of bacteriophage-λ (Table 2.1) were prepared from a 500 ml culture of *E. coli* LE392 (Table 2.1) by inoculating 1 ml of an overnight culture of *E. coli* LE392 grown in LB medium containing 10 mM MgSO₄ into each of four 1 litre conical flasks containing 125 ml of the same medium. These cultures were grown with vigorous shaking at 37°C for 2 h, at which point 10⁹ pfu of bacteriophage-λ were added. Growth was continued until significant lysis was observed (4-6 h). At this point, lysis of *E. coli* was brought to completion by the addition to each flask of 2% (v/v) chloroform followed by a further incubation for 10 min at 37°C with shaking. Solid NaCl was then added to a final concentration of 1 M and cultures placed on ice for at least 1 h. Debris was removed by centrifugation (8 000 g, 4°C, 10 min). The supernatant was collected prior to the purification of bacteriophage-λ DNA.

Streptococcus salivarius ATCC 25975 (Table 2.1) was grown anaerobically at 37°C in semi-defined medium (SDM) (Appendix 1; Jacques, 1983) or on solid Columbia blood agar plates supplemented with 5% (v/v) defibrinated horse blood.

Cultures of *Streptococcus gordonii* LGR2 (Table 2.1) were grown in SDM or Todd Hewitt broth supplemented with 1% (v/v) inactivated horse serum (THBS) and 0.6% (w/v) yeast extract in screw capped tubes containing 20 ml of medium. Individual colonies were grown on Brain Heart Infusion (BHI) agar plates containing 3% (w/v) raffinose. Erythromycin (40 µg ml⁻¹) was added where appropriate. Both agar plates and liquid cultures were incubated anaerobically (5% [v/v] CO₂ / 95% [v/v] N₂) without shaking at 37°C for 1-3 days.

2.4 DNA MANIPULATIONS

2.4.1 Purification of Bacteriophage-λ DNA

Bacteriophage- λ DNA was isolated from 500 ml liquid lysates by the glycerol step gradient method of Silhavy *et al.* (1984). Solid polyethylene glycol (PEG 8000) was added to a final concentration of 10% (w/v) to the centrifuged lysate and incubated on ice for 60 min to allow precipitation of the bacteriophage- λ . The pellet was collected by centrifugation (10 000 g, 4°C, 10min), and resuspended in 10 ml TM buffer (50 mM Tris-HCl pH 7.8, 10 mM MgSO₄). This PEG-bacteriophage- λ DNA suspension was extracted with an equal volume of chloroform and the aqueous phase recovered by centrifugation (4 000 g, 4°C, 15 min). The remaining PEG-chloroform phase and interface was re-extracted with an additional 5 ml of TM buffer and the aqueous phases combined. A glycerol step-gradient was prepared in a 30 ml polycarbonate centrifuge tube by overlaying 6 ml of 40% (v/v) glycerol with 6 ml of 5% (v/v) glycerol before careful addition of the aqueous bacteriophage- λ DNA suspension. This was then centrifuged (125 00 g, 4°C, 60 min). The pellet was resuspended in 1 ml of TM buffer and digested with 5 μ g DNase ml⁻¹ and 1 μ g RNase A ml⁻¹ at 37°C for 30 min. STEP buffer (50 mM Tris-HCl pH7.5 containing 40 mM EDTA, 0.5% [w/v] SDS and 1 mg proteinase K ml⁻¹) was added to a final 0.2 volume and incubated at 56°C for 15 min. The DNA solution was then sequentially extracted with equal volumes of Tris-HCl-saturated phenol (100 mM Tris pH 8.0), phenol:chloroform:isoamylalcohol (25:24:1 [v/v/v]) and chloroform:isoamylalcohol (24:1 [v/v]) with centrifugation (4000 g, 18-20°C, 5 min) to separate the aqueous phase. The DNA was precipitated with 2 volumes of ethanol, spooled onto a Pasteur pipette and resuspended in 500 μ l TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). The resulting DNA suspension was drop-dialysed on a 0.025 μ m Millipore Type VS filter against TE buffer before use.

2.4.2 Rapid Purification of Phagemid DNA

Phagemid DNA was isolated from *E. coli* by a modification of the alkaline lysis method described by Silhavy *et al.* (1984). A single colony of an appropriate strain of *E. coli* was inoculated into 5 ml of LB medium (2.3) containing an appropriate antibiotic and shaken overnight at 37°C. A 1.5 ml sample of this culture was centrifuged (6 000g, 4°C, 5min) and the cell pellet completely resuspended in 80 μ l ice-cold (4°C) solution I (25 mM Tris-HCl pH 8.0 containing 50 mM glucose, 10 mM EDTA and 5 mg lysozyme ml⁻¹) before the addition and mixing by gentle inversion of 160 μ l of solution II (0.2 M NaOH containing 1% [w/v] SDS). After storing on ice for 5 min, 250 μ l of ice-cold (4°C) solution III was added by vigorously inversion of the tube before being left on ice for 5 min. Solution III was prepared by adding 60 ml of 5 M potassium acetate to 11.5 ml glacial acetic acid and 28.5 ml H₂O.

Cellular debris was removed by centrifugation (12 000 g, 4°C, 10 min) and the supernatant retained. Lipids were extracted from the supernatant by the addition of an equal volume of chloroform:isoamylalcohol (24:1 [v/v]) and centrifugation (12 000 g, 4°C, 10 min). The aqueous layer was removed and the DNA precipitated by the addition of 250 µl of 7.5 M ammonium acetate and 750 µl isopropanol. After incubating at -70°C for 15 min, the pellet of DNA was recovered by centrifugation (12 000 g, 4°C, 5 min), rinsed with 70% (v/v) cold (-20°C) ethanol, dried under vacuum and then resuspended in 50-100 µl TE buffer (2.4.1).

Higher purity phagemid DNA required for certain cloning, mutagenic and sequence-screening procedures was purified from *E. coli* cultures with the Wizard™ Plus Miniprep DNA Purification Kit according to the manufacturer's protocol (Promega, WI, USA).

2.4.3 Small Scale Purification of Plasmid DNA from *Streptococcus gordonii*

Plasmid DNA was isolated from *Streptococcus gordonii* LGR2 following growth at 37°C for 16-40 h in 10ml THBS containing an appropriate antibiotic (2.3). Cells were harvested (4 000 g, 18-22°C, 10 min) and resuspended in 200 µl of 10mM sodium phosphate buffer pH 6.5 containing 10mM MgSO₄ and 1mM CaCl₂. The cells were then incubated at 60°C for 15 min with 50 µg mutanolysin before the addition of 20 µl of 10% (w/v) SDS and a further incubation at 18-22°C for 10min. Cells were then lysed at 4°C for 5min following the addition of 100µl of solution II (2.4.2) containing 5 µl of a 10mg ml⁻¹ solution of RNase in 50 mM Tris-HCl pH 7.5. Cellular debris was precipitated with 300 µl of solution III (2.4.2) at 4°C for 15min before centrifugation (13 000 g, 4°C, 15 min). The supernatant was then extracted with an equal volume of Tris-HCl-saturated phenol (2.4.1), followed twice with an equal volume of chloroform:isoamylalcohol (24:1; v/v). The plasmid DNA was finally precipitated at -20°C for 15 min by the addition of 2.5x the volume of ethanol, and recovered by centrifugation (13 000 g, 4°C, 15 min). The pelleted DNA was dried under vacuum and resuspended in 50 µl TE buffer (2.4.1).

2.4.4 Large Scale Purification of Phagemid DNA

Phagemid DNA was purified from *E. coli* on a large scale according to the method described by Sambrook *et al.* (1989). Essentially, *E. coli* was grown overnight (16 h) at 37°C with vigorous aeration in 500 ml LB medium (2.3) with the appropriate antibiotic. Cells were harvested by centrifugation (8 000 g, 4°C, 15 min) and resuspended in 10 ml solution I (2.4.2). The cells were then kept at room temperature (18-20°C) for 5 min before 20 ml of solution II (2.4.2) was added and the cell suspension incubated on ice for a further 10 min.

Solution III (15 ml; 2.4.2) was then added and bacterial debris removed by centrifugation (12 000 g, 4°C, 30 min). The supernatant was carefully removed and the DNA precipitated with 2 volumes of ethanol. The DNA was recovered by centrifugation (12 000 g, 18-22°C, 30 min), vacuum dried and resuspended in 16 ml TE buffer (2.4.1). This solution was transferred to an ultracentrifuge tube before cesium chloride was added to a final concentration of 1 g ml⁻¹ together with 1.6 ml of ethidium bromide solution (10 mg ml⁻¹). The ultracentrifuge tube was filled with light paraffin oil, sealed and centrifuged (200 000 g, 20°C, 36 h). The lower band consisting of closed circular phagemid DNA was visualized under UV light and removed with a hypodermic needle and syringe. The ethidium bromide was extracted several times with equal volumes of H₂O-saturated 1-butanol and the resulting aqueous phase dialysed against three changes of TE buffer (2.4.1). The DNA was precipitated by the addition of 0.1 volume of 3 M sodium acetate pH 4.8 and 2.5 volumes of ethanol. The precipitated DNA was stored in ethanol at 4°C and recovered by centrifugation (12 000 g, 22°C, 10 min) when required.

2.4.5 Preparation of Single-Stranded Phagemid DNA

Phagemid single-stranded DNA (ssDNA) and uricilated ssDNA was generated in *E. coli* stains NM522 and CJ236 respectively (Table 2.1) carrying appropriate phagemids. A 10 ml culture of '2xYT' medium (2.3) containing 500 µg ampicillin ml⁻¹ and 250 ng uridine ml⁻¹ in the case of *E. coli* CJ236, was inoculated with 200 µl of an overnight (16 h) culture of the appropriate strain of *E. coli* carrying the phagemid of interest and grown with shaking at 37°C to mid-exponential phase. A 2 ml portion of this culture was then transferred to a fresh flask and 10⁸ pfu bacteriophage M13KO7 ml⁻¹ added (10µl; Table 2.1) and incubated with shaking at 37°C. After 60 min, a 400 µl sample was transferred to 10 ml of fresh '2xYT' medium supplemented with 500 µg ampicillin ml⁻¹ and 70 µg kanamycin ml⁻¹ as well as 250 ng uridine ml⁻¹ in the case of *E. coli* CJ236. Cultures were incubated with vigorous aeration at 37°C for 16 h to allow for cell lysis. Cell debris was removed by centrifugation (12 000 g, 4°C, 10 min) and 1 ml 25% (w/v) PEG (8000) in 3 M NaCl added to the supernatant prior to it being stored on ice for 60 min. The precipitated bacteriophage were recovered by centrifugation (10 000 g, 4°C, 30 min) and resuspended in 400 µl of 20mM Tris-HCl pH 7.6 containing 20 mM NaCl and 1 mM EDTA. The suspension was extracted twice with Tris-HCl-saturated phenol (2.4.1) and twice with chloroform:isoamylalcohol (24:1 [v/v]) to release phagemid ssDNA from the bacteriophage by removing the protein coat. The ssDNA was precipitated on ice for 30 min with 0.1 volume of 3 M sodium acetate pH 5.2 and 2 volumes of ethanol before being recovered by centrifugation (12 000 g, 18-22°C, 10 min). The ssDNA was dried under

vacuum and resuspended in 50 μl of TE buffer pH 7.6 (2.4.1) prior to verification by agarose gel electrophoresis (see 2.4.7 below).

2.4.6 Spectrophotometric Determination of Nucleic Acid Concentration

The concentration and purity of 5 μl samples of nucleic acid were determined spectrophotometrically at A_{260} and A_{280} nm using a Pharmacia GeneQuant DNA/RNA Calculator (Sambrook *et al.*, 1989). An A_{260} of 1.000 corresponded to approximately 50 μg DNA ml^{-1} for dsDNA and 40 μg DNA ml^{-1} for 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 Electrophoresis of DNA on Agarose Gels

DNA was separated on 0.8% (w/v) or 1.5% (w/v) agarose gels prepared in TBE buffer (89 mM Tris pH 8.0 containing 89 mM boric acid and 2 mM EDTA) using either the Hoefer 50 ml 'mini-gel' or 150 ml 'maxi-gel' system (Hoefer Scientific Instruments, Ca, USA). Electrophoresis was carried out in TBE buffer at 80-120 V for 2-6 h or 25 V overnight (16 h). Samples were diluted to 20 μl in TE buffer pH 8.0 (2.4.1) and 3 μl loading dye (30% [v/v] glycerol containing 0.25% [w/v] bromophenol blue and 0.25% [w/v] xylene cyanol FF) was added. Following electrophoresis, the agarose gels were stained in a dilute ethidium bromide solution (5 μg ml^{-1}) for 1h. Agarose gels were photographed on a UV transilluminator (Novex Australia Pty. Ltd., Sydney, Australia) with a polaroid camera fitted with a Kodak Wratten 23A orange filter using type 667 Polaroid film (Polaroid Australia Pty. Ltd., Sydney, Australia).

2.4.8 Recovery of DNA from Agarose Gels

DNA was recovered from agarose gels using a modification of the method of Weichenban (1991). Whitman 3M filter paper was wetted in Repel-silane and allowed to dry in a fume-hood. The treated filter paper was rinsed twice with 250 ml aliquots of 50 mM Tris-HCl pH 8.0 containing 1 mM EDTA and once in 250 ml of deionized H_2O before drying in a 37°C oven. The dried paper was cut into squares (2 cm x 2 cm), which were folded to fit slices of agarose gel that had been removed after electrophoretic separation of DNA samples and contained the DNA band of interest (2.4.7). These agarose-filled envelopes were inserted into a microfuge tube which had had its base punctured and was inserted into a second microfuge tube. The DNA was recovered as a solution in the lower microfuge tube by centrifugation (6 000 g, 18-22°C, 2 min). The DNA was extracted with water-saturated

butanol to remove ethidium bromide before precipitating with ethanol and storing as described above (2.4.4).

2.4.9 Digestion of DNA with Restriction Endonucleases

Restriction endonuclease digestion of DNA was routinely carried out in volumes of 30-200 μ l, using the restriction buffers and conditions recommended by the manufacturer. Incubation was carried out for 1-3 h at the optimal temperature required for the particular restriction endonuclease. Heating at 65°C for 20 min was routinely used to inactivate the restriction endonuclease prior to agarose gel electrophoretic analysis of a sample of the digested DNA (2.4.7).

2.4.10 Dephosphorylation of DNA

Self-ligation of vector DNA was prevented by dephosphorylation using calf-intestinal alkaline phosphatase. The heat inactivated and precipitated restriction digested DNA (2.4.9) was resuspended in 50 μ l of 1x alkaline phosphatase buffer supplied by the manufacturer. Alkaline phosphatase (1-2 U [μ g DNA]⁻¹) was added and the solution incubated at 37°C for 60 min. The reaction was stopped by the addition of 5 mM EDTA and further incubation at 75°C for 10min. The DNA was recovered by precipitation with 2 volumes of ethanol (2.4.3).

2.4.11 DNA Ligation

DNA that was to be ligated into plasmid or phagemid vectors was precipitated with ethanol (2.4.3) and resuspended in a total volume of 20-50 μ l of 1x ligation buffer (supplied by the manufacturer of the T4 DNA ligase). The amount of insert DNA added was 3 – 5 x the amount of vector DNA as determined by visualization on agarose gels and/or by spectrophotometric measurement at A₂₆₀ and A₂₈₀ nm (2.4.6). The total amount of DNA in the ligation was usually between 0.5 – 2.0 μ g. T4 DNA ligase (25 U, 1 μ l) was added to the solution and incubated at 4°C for 16 h after which time another 25 U (1 μ l) of T4 DNA ligase was added and the incubation continued for 2 h at 18-22°C. Samples with high salt concentrations were precipitated with ethanol (2.4.3) before further use.

2.4.12 Exonuclease Digestion of DNA

Phagemid DNA purified by equilibrium centrifugation in CsCl-ethidium bromide gradients (2.4.4) was linearized with the restriction enzyme *Xba*I and then digested with *Apa*I (2.4.9) prior to the production of nested sets of deletions.

Precipitated DNA (30 µg) was dissolved at 37°C in a 60 µl 66 mM Tris-HCl pH 8.0 containing 6.6 mM MgCl₂. At zero time 1 000 U of exonuclease III was added and 5 µl samples subsequently removed at 45 s intervals and added to 15 µl S1 reaction mix on ice. S1 reaction mix was prepared by mixing 172 µl H₂O with 60 U nuclease S1 and 27 µl 10x S1 buffer. S1 buffer, itself, was prepared by mixing 1.1 ml 3 M potassium acetate pH 4.5 with 5 ml 5M NaCl, 5 ml glycerol and 20 µl 1M ZnSO₄. Following the collection of the series of samples at 45 s intervals in the S1 reaction mix, the samples were incubated for a further 30 min at 30°C. This reaction was stopped by the addition of 2 µl of S1 stop mixture (3 mM Tris-HCl pH 8.0 containing 50 mM EDTA) and by incubating the samples at 70°C for 10 min. The ‘fill-in’ of the single stranded ends resulting from the exonuclease III treatment was achieved by the addition to each tube of 3 µl of a 50 µl Klenow reaction mix containing 10 mM Tris-HCl pH 7.6, 200 mM MgCl₂ and 10 U Klenow DNA Polymerase I and incubating at 37°C for 5 min before the subsequent addition to each tube of 4 µl of a solution containing each of the four deoxynucleotide phosphates (dNTP) each at a concentration of 500 µM and incubating for a further 30 min at 18-22°C. Aliquots (5 µl) of each time point sample were analyzed by agarose gel electrophoresis (2.4.7). The remaining 25 µl of each sample was precipitated with 2.5 volumes of ethanol and dried (2.4.4). Samples containing DNA fragments of the required size were reconstituted for ligation into phagemid vectors (Table 2.1; 2.4.11) and subsequently transformed (2.4.14; 2.4.15 below).

2.4.13 Preparation of Competent *E. coli*

Strains of *E. coli* that were to be transformed by electroporation were made competent using a method derived from the instruction manual for the ‘Gene-Pulser’ electroporation unit (Bio-Rad, Sydney, Australia). A cultures of a relevant strain of *E. coli* was grown in 1litre of LB medium (2.3) at 37°C to an A₆₀₀ of 0.500, at which time the culture was harvested by centrifugation (4 000 g, 4°C, 15 min) and the cells resuspended on ice at 4°C in 200 ml 10% (v/v) glycerol for 20 min. This process of centrifugation and incubation was repeated before the cells were finally harvested by centrifugation (4 000 g, 4°C, 15 min) and resuspended in 2 ml ice-cold 10% (v/v) glycerol. The final suspension of *E. coli* cells was dispensed into 50 µl aliquots and stored frozen at -70°C until required.

2.4.14 Transformation of *E. coli* by Electroporation

An aliquot of competent *E. coli* cells (60 µl) that had been prepared for electroporation (2.4.13) was thawed on ice, 5 µl of DNA (0.05-0.2 µg) added and the mixture of cells and

DNA transferred to a ice-cold electroporation cuvette with a 0.1 cm electrode gap. The ‘Gene Pulser’ (BioRad, Sydney, Australia) was set at 25 μ F and the ‘Pulse Controller’ (BioRad, Sydney, Australia) to 200 . A pulse of 1.6 kV was applied to the cuvette and ice cold 2xYT medium (1 ml; 2.3) immediately added and the cell suspension transferred to a microfuge tube and incubated at 37°C with shaking (150 rpm) for 1 h before the bacteria were plated on selective medium.

2.4.15 Natural Transformation of *S. gordonii*

S. gordonii LGR2 (Table 2.1) was transformed according to the protocol of Drucker *et al.* (1989).

A stock filtrate containing competence factor was made by inoculating 10 ml of Todd Hewitt Broth (THB) containing 5% (v/v) heat-inactivated horse serum with 0.5 ml of an overnight (16 h) culture of *S. gordonii* and incubating at 37°C for 75 min. Cells were removed by centrifugation (20 000 g, 4°C, 15 min) and the supernatant filtered through a 0.2 μ m Sartorial minisart filter prior to being dispensed in 400 μ l aliquots and stored at –20°C as a source of competence factor.

In order to transform *S. gordonii*, 100 μ l of an overnight (16 h) culture of *S. gordonii* LGR2 was inoculated into 1.9 ml of THBS (2.3) and incubated at 37°C for 60 min. A 40 μ l sample of this culture was then inoculated at 37°C into 400 μ l filtrate containing competence factor to which 1-2 μ l of appropriate plasmid DNA had been added. After incubating at 37°C for 2 h to allow adsorption and uptake of DNA, cells were plated onto appropriate selective media.

2.4.16 DNA Sequencing

Sequence determinations were carried out using cesium chloride purified or WizardTM Plus Minipreps purified dsDNA (2.4.2; 2.4.3; 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 the supplied T7, T3 or custom-synthesized oligonucleotide primers and [³⁵S]-dATP (2.1).

Sequencing reactions were run using a Hoefer Poker Face I Nucleic Acid Sequencer (Hoefer Scientific Instruments, California, U.S.A.). Sequencing gels [8% (w/v) acrylamide, 28% (w/v) urea] were pre-run at 110 W for at least one hour. Three loadings were applied to the gel, at 0 h, 1.5 h, and 3.5 h, with a total running time of approximately 5.5 h. After loading, the gel was run at 10 W until the loading dye moved into the gel (5-10 min), 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 FRUCTOSYLTRANSFERASE ACTIVITIES AND PROTEIN MANIPULATIONS

2.5.1 Detection of Sucrase Activity

Sucrase activity was detected using a qualitative microtitre reducing-sugar test based on the method described by Aduse-Opoku *et al.* (1989). A loopful of cells or 50 µl of culture were suspended in 100 µl of 50 mM sodium phosphate buffer pH 6.5 containing 0.1% SDS and incubated at 18-22°C for 15 min. Following this incubation, 100 µl of sucrose buffer consisting of 50 mM sodium phosphate pH 6.5 containing 1% (v/v) Triton X-100, 1% (w/v) sucrose and 0.01% (w/v) thiomersal was added and the incubation continued at 37°C for 60 min. Subsequently, 100 µl of developing solution consisting of 1 M NaOH containing 0.1% (w/v) triphenyltetrazolium chloride was added and the sample allowed to incubate at 37°C for a further 15 min. The development of a red colour indicated the presence of released reducing sugars.

2.5.2 Assay for Fructosyltransferase Activity

Ftf activity was expressed as the polymer-forming activity determined by the formation of radioactively labelled fructan at 37°C (Jacques & Wittenberger, 1981). The reaction was initiated by the addition of 0.5 ml of 40 mM [U-¹⁴C-fructosyl]-labelled sucrose (1.8 kBq ml⁻¹) into a 1.5 ml of reaction mixture containing 100 mM potassium phosphate buffer pH 6.0, 10 mM NaF, 100 µM histidine, 1 mM CaCl₂ and an appropriate amount of Ftf. At the end of 60 min the reaction was stopped by the addition of 6 ml of ethanol to the reaction tube. The precipitated fructan was then filtered through a 2.5-cm Whatman GAF/B glass fibre filter under vacuum. Reaction tubes and filters containing the trapped precipitated fructan were each washed three times with 5 ml of 75 % (v/v) ethanol. The dried filters were suspended in Ultima-Gold[™] scintillation solution and counted in a liquid scintillation spectrometer (Beckman LS 9000, Beckman Instruments, Sydney, Australia). One unit of enzyme activity (U) was defined as the amount of Ftf that catalyzed the incorporation of 1 µmol of the fructose moiety of sucrose into 75 % (v/v) ethanol-insoluble polysaccharide per min.

2.5.3 SDS-PAGE Gel Electrophoresis

SDS-PAGE gels were run on a Hoefer SE 600 system (Pharmacia Biotech, Melbourne,

Australia) or a Bio-Rad Protean II xi cell (Bio-Rad, Sydney, Australia) according to the method of Laemmli (1970). Samples of proteins (20-200 μ l) were solubilized at 100°C in a third the volume of SDS sample buffer (200mM Tris, 40% [v/v] glycerol, 4% [w/v] SDS, 20% [v/v] β -mercaptoethanol, 0.001% [w/v] bromophenol blue) for 5 min before applying to the gel. Broad Range SDS-PAGE molecular weight markers (1 μ l) were similarly treated prior to application to the gel. For the Hoefer SE 600 system, the gels were run at 15 mA until the tracking dye passed through the stacking gel when the current was increased to 30 mA, or alternatively, gels were run at 60 V overnight for 16 h. For the Bio-Rad Protean II xi cell, gels were run at 200 V until the loading dye moved out of the gel (~1 h). The gels were then fixed in 10% (w/v) trichloroacetic acid for 30 min, rinsed twice with deionized water and stained with 'Gradipure' stain overnight, before destaining in 10% (v/v) methanol. In some cases, the gels were silver-stained or activity-stained as described below (2.5.4; 2.5.5).

2.5.4 Detection of Proteins on SDS-PAGE with Silver Staining

Silver staining was carried out according to the method described by Swain and Ross (1995). After electrophoresis, gels were washed in 18.2 M Ω deionized water for 5 min before being fixed in 200 ml of ethanol/acetic acid/water (40:10:50, by volume) for 30 min. After fixing, the gels were shaken in ethanol:acetic acid:water (10:10:180 [v/v/v]) for 1-16 h, washed in 18.2 M Ω water for 5 min, soaked in 0.5 M sodium acetate containing 1% (v/v) glutaraldehyde for 30 min, washed in 18.2 M Ω water 3 x 10 min and stained in 150 ml of silver-staining solution for 30 min (2 ml of concentrated ammonia added to 28 ml of 0.1 M NaOH followed by slow addition of 5 ml 20% [w/v] AgNO₃ with mixing and then 150 ml of 18.2 M Ω water). After a further 4 x 10 min washes in 18.2 M Ω water, the gels were developed in 200 ml of 0.05% (w/v) citric acid solution containing 0.1% (v/v) formaldehyde until the protein bands could be seen (1-10 min). Acetic acid (5%, v/v) was then added to stop the colour development.

2.5.5 Detection of Fructosyltransferase Activity in SDS-PAGE

Ftf activity in SDS-PAGE gels was determined by a periodic acid Schiff's (PAS) reaction (Pitty *et al.*, 1989). After electrophoresis (2.5.3), SDS-PAGE gels were rinsed briefly in 18.2 M Ω water and then incubated at 37°C for 16h in 50 mM sodium phosphate buffer pH 6.5 containing 1% (w/v) sucrose, 1% (v/v) Triton X-100, 1mM CaCl₂ and 0.01% (w/v) thiomersal. After rinsing the gels in 18.2 M Ω water they were fixed in 75% (v/v) ethanol for 30 min and then incubated for 30 min in a solution containing 0.7% (w/v) periodic acid and

5% (v/v) acetic acid. After a further 3 x 15 min washes in a 5% (v/v) acetic acid solution containing 0.2% (w/v) sodium metabisulfate, the gels were incubated in Schiff's reagent for 60 min and then rinsed twice for 15 min in a 40% (v/v) methanol solution containing 5% (v/v) acetic acid and 0.5% (w/v) potassium metabisulfate before destaining by repeated washes in a 20% (v/v) methanol solution.

2.5.6 Electroblothing of Proteins from SDS-PAGE onto PVDF Membranes

Electrotransfer of proteins from SDS-PAGE gels to 0.2 μ m polyvinylidene difluoride (PVDF) membranes was carried out on a Hoeffer TE Series Transphor unit (Pharmacia Biotech, Melbourne, Australia) according to the method of Speicher (1997). PVDF membranes were prepared by soaking pre-cut sheets in methanol for 1 min, 18.2 M Ω deionized water for 5 min and then transfer-buffer containing 10 mM Tris base, 100 mM glycine, 10% (v/v) methanol and 100 μ M thioglycolate. SDS-PAGE gels were rinsed twice in 18.2 M Ω water before assembly into the Hoeffer Transfor unit containing transfer buffer and electroblotted at a constant current of 250 mA. The PVDF membranes were stained for 15 s with amido black reagent containing 0.5% (w/v) amido black, 25% (v/v) isopropanol and 10% (v/v) acetic acid, before being destained with 6x changes of 18.2 M Ω water for 5 min and air-drying at room temperature (18 – 20⁰C).

2.6 DATA PROCESSING AND BIOINFORMATIC ANALYSES

Sequenced DNA was analyzed using the IBI Pustell sequence analysis software version 2.03 (IBI corporation, Toronto, Ontario, Canada). Database searches were carried out using the Fasta program (Pearson and Lipman, 1988) and/or the Blast program (Altschul *et al.*, 1990). Multiple sequence alignments and phylogenetic analyses used the Clustal V program (Higgins *et al.*, 1992) and the PHYLIP (Phylogeny Interference Package) version 3.5 package of software (Felsenstein, 1989) respectively, while the Terminator program (Brendel and Trifonov, 1984) in the Genetics Computer Group Inc (GCG) package of analytical software was used to search for a factor-independent transcription terminator. Fasta, Blast, ClustalV and GCG were accessed through the Australian National Genomic Information Service (ANGIS; <http://www.angis.org.au>).

CHAPTER 3

ANALYSIS OF THE REGION FLANKING THE *ftf* GENE IN BACTERIOPHAGE λ A-39¹

3.1 INTRODUCTION

Streptococcus salivarius ATCC25975 (Table 2.1) synthesizes a Ca^{2+} -dependent Ftf which is initially cell-associated (Garszczynski and Edwards, 1973; Jacques and Wittenberger, 1981) but released on exposure to its substrate, sucrose (Milward and Jacques, 1990). To better understand the structure and hence the interaction of the Ftf protein with the cell, a gene bank of chromosomal DNA from *S. salivarius* was constructed with 5-15 kbp *Sau*3A fragments cloned into bacteriophage λ L47.1 (Pitty *et al.*, 1989). Sequencing of polymer-forming λ -recombinants resulted in the isolation of λ A-39 which contained a 8.2 kbp DNA insert coding for a Ca^{2+} -dependent Ftf (Pitty *et al.*, 1989). In order to isolate and map the *ftf* gene, λ A-39 was subcloned and the resulting inserts sequenced.

3.2 METHODS

3.2.1 Sequencing of DNA

DNA from bacteriophage λ A-39 (Table 2.1) that was to be sequenced was first subcloned into either pIBI30 or pIBI31 (Table 2.1) and propagated in *E. coli* strain JM109 or NM522 (Table 2.1). Sequencing reactions were carried out on ssDNA or CsCl purified dsDNA as described in Sections 2.4.5 and 2.4.16. The DNA sequence was confirmed in both directions. The nucleotide sequence data were submitted to GenBank and assigned the accession numbers L07793 for the region containing *orf2* and *orf3* and L07794 for the region containing *orf4*.

3.2.2 DNA Sequence Analysis.

DNA sequences were assembled and open reading frames (ORFs) detected using the IBI-Pustell sequence analysis software version 2.03. Database searches were carried out using the "Fasta" program (Pearson & Lipman, 1988) while multiple sequence alignments and phylogenetic comparisons used the "ClustalV" program (Higgins *et al.*, 1992). This program makes use of the neighbour-joining method of Saitou & Nei (1987) to deduce phylogenetic trees. Fasta and ClustalV were accessed through ANGIS (2.6).

¹ The results presented in this chapter have already been published – Giffard *et al.*, *J. Gen. Microbiol.* 139: 913-920, 1993.

3.3 RESULTS AND DISCUSSION

3.3.1 Physical Map of *Streptococcus salivarius* ATCC25975 Chromosomal DNA in λ A-39

Mapping of the chromosomal DNA insert was performed by means of double restriction digests in relation to known sites in bacteriophage- λ . Subclones were then constructed based on this map (Fig.3.1; Table 2.1).

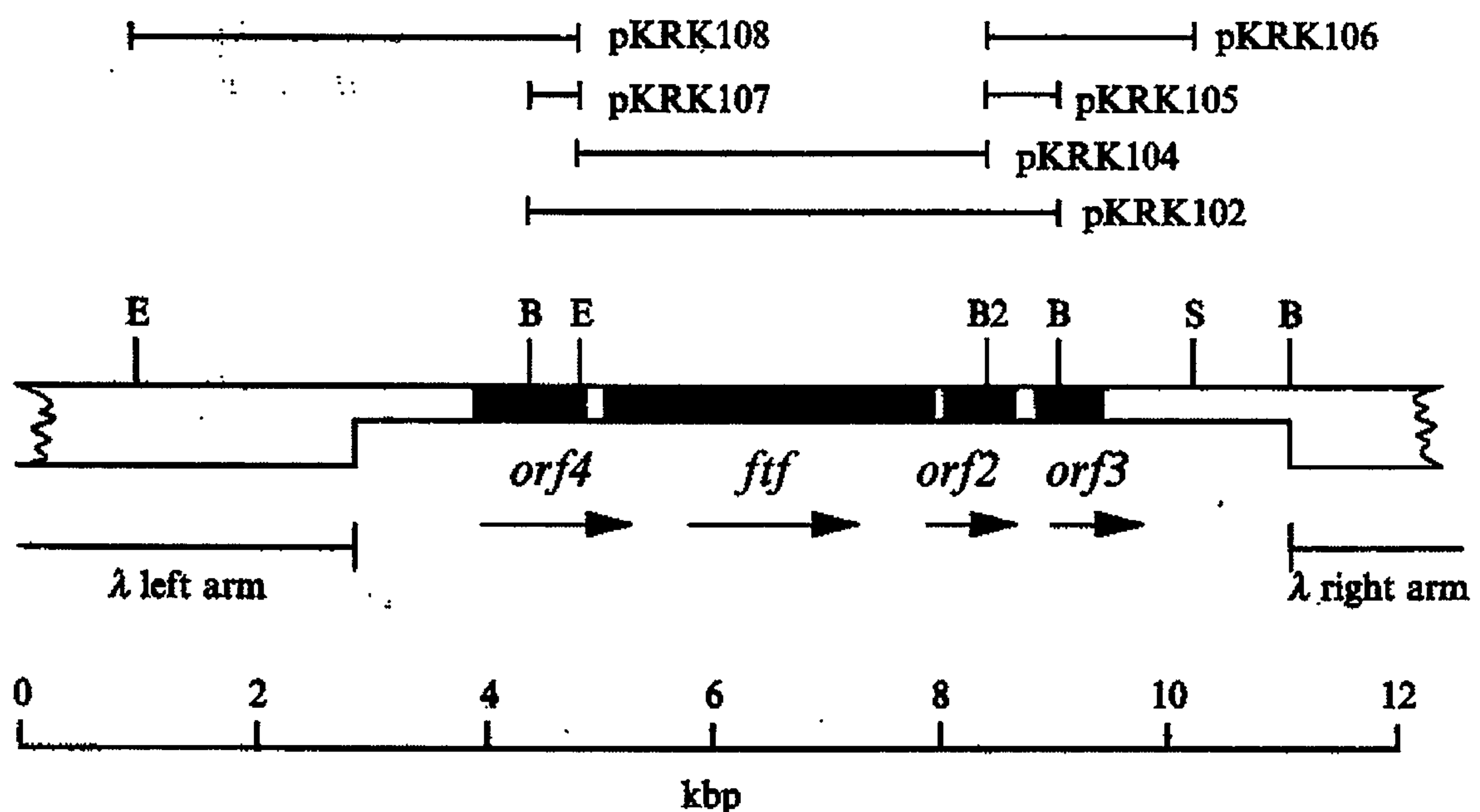


Figure 3.1. Physical map of the *ftf* locus of *S. salivarius* ATCC 259575. The fragments cloned into pIBI30/31 to form the pKRK series of plasmids are shown above the map. Coding regions are indicated by the filled regions. E, *EcoRI*; B, *BamHI*; B2, *BglIII*; S, *SacI*.

The first subclone constructed was the phagemid pKRK102, which contained the 4.6 kbp *BamHI* fragment from the *S. salivarius* chromosomal insert in λ A-39. This phagemid conferred a strong *Ftf*⁺ phenotype on the *E. coli* host in an assay for reducing sugars (2.5.1). While this indicated that the *ftf* gene was expressed by this phagemid in *E. coli*, the phagemid construct proved impossible to maintain (Table 3.1). The *E. coli* host died after a few rounds of subculturing and propagation in liquid media resulted in a loss of the phagemid (when growth was present). The deleterious effects of pKRK102 was initially thought to be due to overproduction of the *Ftf* protein which could feasibly cause blockage of the host protein export machinery. However, on further subcloning the phagemid pKRK104, containing the 3.6 kbp

EcoRI-BglII fragment from the *S. salivarius* chromosomal insert in λA-39, was isolated from a stable Ftf⁺ host as was the phagemid pKRK107 containing the 0.5 kbp *BamHI-EcoRI* fragment. In contrast, pKRK105 containing the 0.6 kbp *BglII-EcoRI* DNA insert was isolated from a viable though slow growing host and with a much reduced yield (Table 3.1). Since the DNA inserts in these three phagemids were contiguous in pKRK102, this suggested that the cause of *E. coli* lethality in pKRK102 may not have been due to the overproduction of Ftf, but possibly due to the presence of the DNA or the product expressed by pKRK105.

Table 3.1. Phenotypes conferred on *E. coli* by phagemid subclones

Phagemid	Ftf Activity	Phagemid Yield
pKRK102	+	(±)*
pKRK104	+	+++
pKRK105	-	++
pKRK107	-	+++

* pKRK102 was almost impossible to maintain and had to be prepared from isolated colonies on solid growth medium.

Sequencing of pKRK104 confirmed the presence of the complete *ftf* gene in this phagemid (Chapter 4) as well as the beginning of another ORF, *orf2*, downstream of the *ftf* (Fig. 3.1). The initial deduced amino acid sequence of the partial *orf2* exhibited strong similarity to the *Saccharomyces cerevisiae FUR1* gene product. Complete sequencing of the pKRK105 and pKRK106 containing the 2.6 *BglII-BamHI* fragment (Kwan, 1991), revealed the presence of the entire *orf2* and a third ORF (*orf3*) (Fig. 3.2). The deduced translation of *orf3* showed homology to the ClpP family of proteins. Sequencing of pKRK107 and pKRK108, containing the *EcoRI* – *EcoRI* fragment upstream of the *ftf* gene, showed the presence of *orf4* (Fig. 3.3) which, when translated, revealed high homology with three known transposases. All ORFs were transcribed in the same direction.

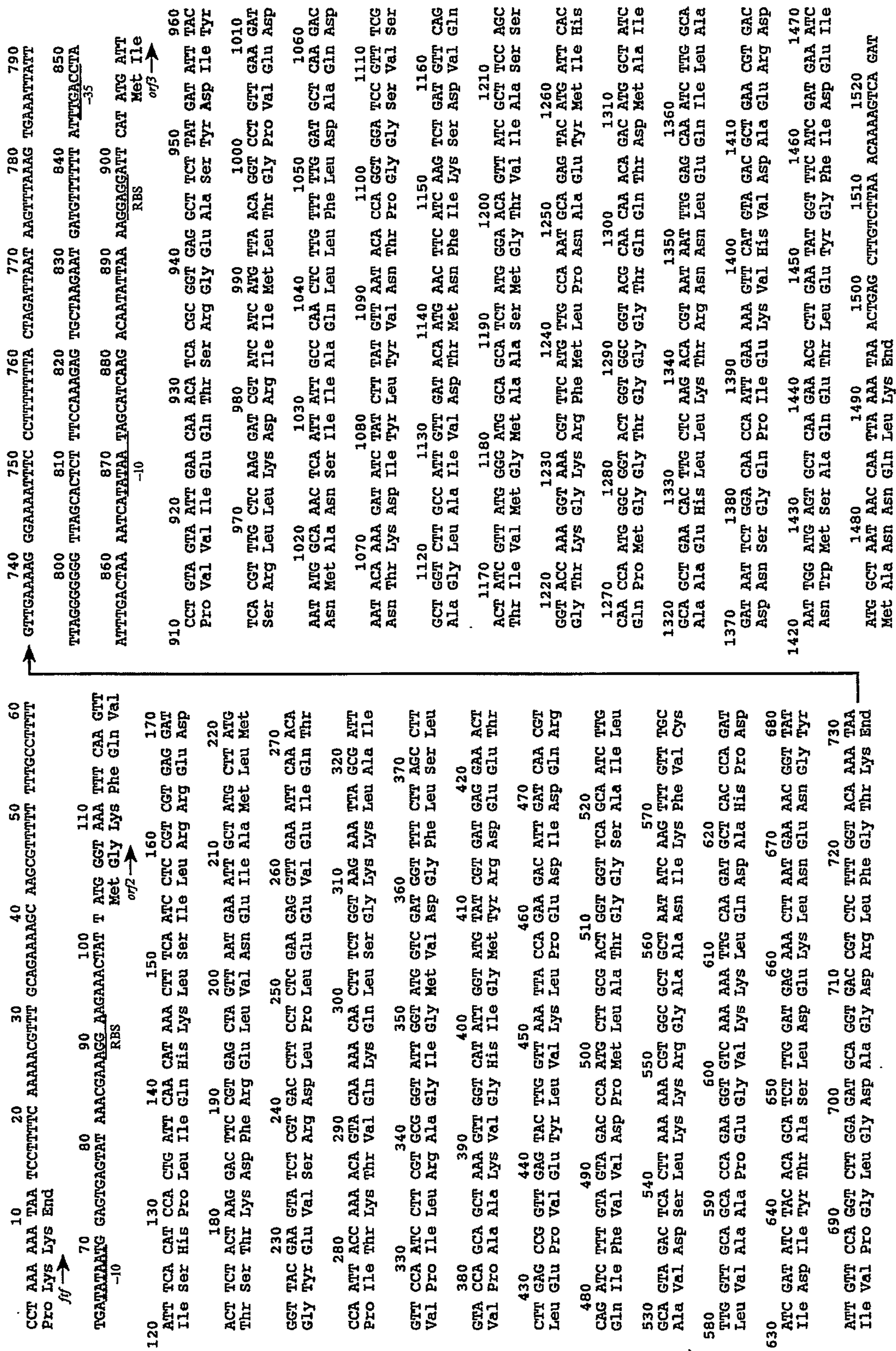


Figure 3.2. Nucleotide sequence of DNA immediately downstream from the *S. salivarius ftf* gene and the deduced amino acid sequence of *orf2*, *orf3* and the 3'-OH end of the *fft* gene. The putative ribosome-binding sites (RBS) are underlined as are the putative promoter sites except for the -35 region upstream of *orf2* where no clear homology to the *E. coli* consensus sequence could be discerned.

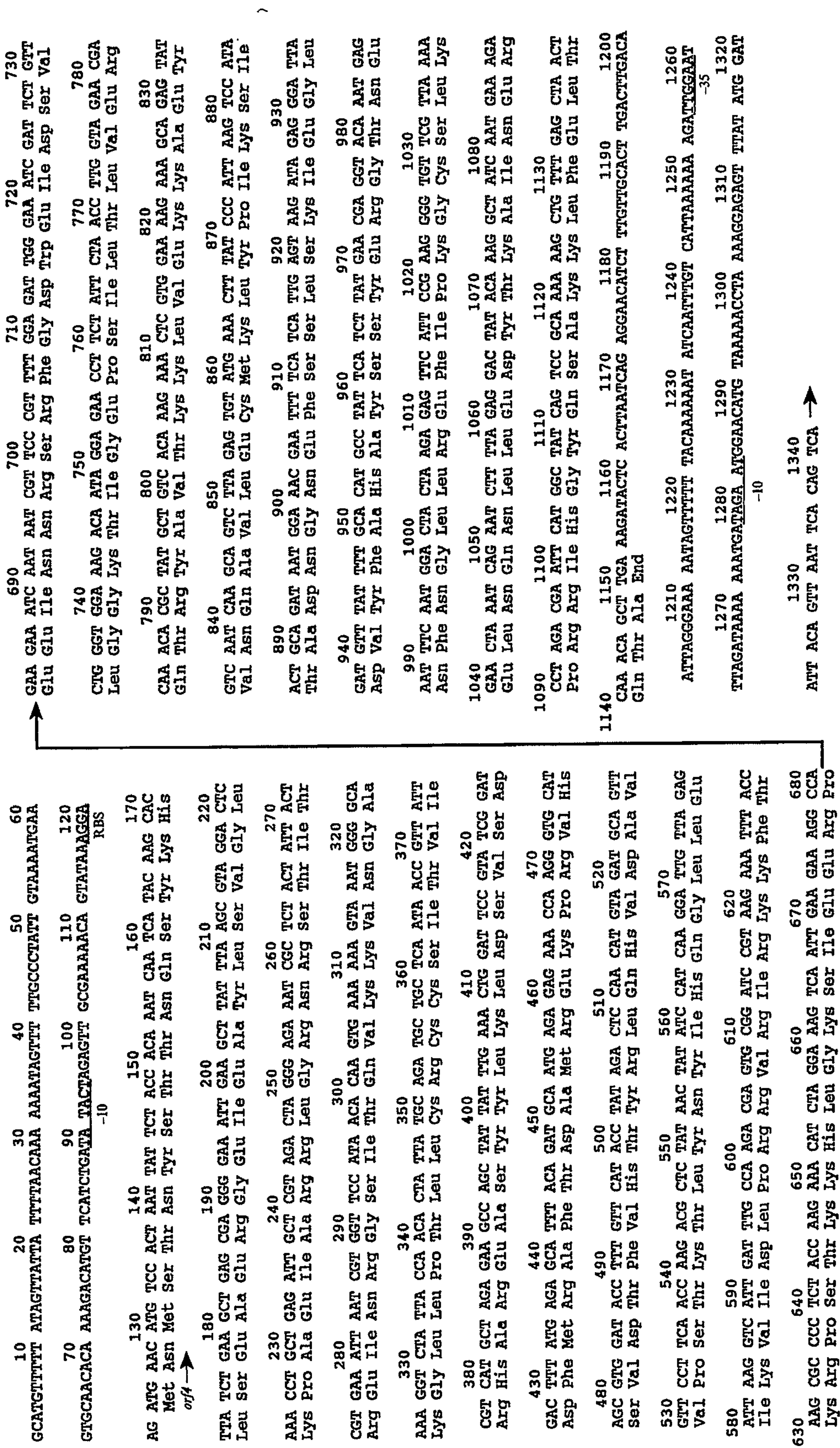


Figure 3.3. Nucleotide sequence of DNA immediately upstream from the *S. salivarius ftf* gene and the deduced amino acid sequence of *orf4*. Putative ribosome-binding sites (RBS) are underlined as are the putative promoter sites, except for the -35 region upstream of *orf4* where no clear homology to the *E. coli* consensus sequence could be discerned.

3.3.2 The ORF Downstream of the *ftf* Gene Codes for a *FUR1* Homologue

The transcription start site of *orf2* lay 90 bp downstream of the *ftf* gene (Fig 3.2). The deduced amino acid sequence of this ORF showed 63% homology with the *FUR1* gene product of *Saccharomyces cerevisiae* which encodes a uracil phosphoribosyltransferase (UPRTase) (Fig 3.4; Kern *et al.*, 1990). The UPRTase catalyses the formation of uridine 5'-monophosphate (UMP) from uracil in the pyrimidine salvage pathway (Musick, 1981). Whilst the activity of the product of *orf2* was not assayed, the 33% identity and 63% conservation of amino acids across the prokaryote/eukaryote boundary appears to indicate that there is selective pressure for conservation of function.

Viability of the *E. coli* hosts harbouring the various chromosomal subclones (Table 3.1) indicated that the lethality conferred on *E. coli* by pKRK102 was due to the presence of *orf2* since interruption of this gene, as in pKRK104, resulted in a non-lethal phagemid. The detrimental effect of the presumptive *S. salivarius* UPRTase on the metabolism of *E. coli* was not self evident since *E. coli* expresses its own UPRTase (Rasmussen *et al.*, 1986). However, overproduction of this enzyme may result in derangement of nucleotide synthesis. Interestingly, it has been reported that over expression of another protein involved in pyrimidine biosynthesis, the *Bacillus subtilis* uracil permease (PyrP), is toxic to both *E. coli* and *B. subtilis* cells (Turner *et al.*, 1994). It appears that the strong homology between the UPRTases of various organisms and tight regulation of the enzymes in this pathway are essential for cell viability.

3.3.3 The ORF Downstream of the Putative *FUR1* Gene Codes for a ClpP Homologue

The transcription start site of *orf3* lay 173 bp downstream of *orf2* (Fig 3.2). The deduced amino acid sequence of *orf3* exhibited similarity to the ClpP family of heat shock polypeptides (Fig 3.5; Kroh and Simon, 1990). The Clp proteases family is composed of two types of subunits; the ClpP which contains the proteolytic active site and the homologous ClpA or ClpX which act as the ATPase regulatory unit (Maurizi, 1992; Gottesman *et al.*, 1993). It has been demonstrated that the ClpA functions as a molecular chaperone directing protein unfolding for degradation by ClpP (Wickner *et al.*, 1994).

In recent years it has been noted that the *clp* homologues are found close to genes encoding surface or extracellular polypeptides in a range of bacteria. This is of particular interest as the *ftf* gene of *Streptococcus mutans* GS-5 is preceded by a ClpB homologue (Gottesman *et al.*, 1990a). The ClpB proteins are sequence homologues to ClpA, however they do not interact with

<i>S. salivarius</i>	MGKQVISHPLI-----QHKLSILRR-----EDSTKDFRE--LVNEIAMLMG-YEVS
<i>Sacch. cerevisiae</i>	MNPLFFLASPFLYLYLIYYPNKGSFVSKPRNLQKMSSEPKNVYLLPQTNQLGLYTI
	*. . . * * * * *
<i>S. salivarius</i>	RD-----LPLEEVEIQTPITKTV-----QQLSGK-----KLAIVPILR
<i>Sacch. cerevisiae</i>	RNKNTTRPDFIFYSDRIIRLLVEEGLNHLPVQKQIVETDTNENFEGVSFMGKICGVSIVR
	*. * * * *
<i>S. salivarius</i>	AGIGMVDGFLSLVPAAKVGHIGMYRDEETLEPVEYLVKLPEIDQRQIFVVDPMLATGGS
<i>Sacch. cerevisiae</i>	AGESMEQGLRDCCRSVRIGKILIQRDEETALPKLFYEKLPEDISERYVFLLDPMLATGGS
	** . * * * * *
<i>S. salivarius</i>	AILAVDSLKKRG--AANIKFVCLVAAPGVKKLQDAHPDIDIYTASLDEKLNENGYIVPG
<i>Sacch. cerevisiae</i>	AIMATEVLIKRGVKPERIYFLNLICSKEGIEKYHAAFPVRIVTGALDRGLDENKYLVPG
	** . * . . * * * * *
<i>S. salivarius</i>	LG DAGDR LFGTK
<i>Sacch. cerevisiae</i>	LGDFGDRYYCV-
	*** *** . .

Figure 3.4. Alignment of deduced amino acid sequence *orf2* of *S. salivarius* with the URPTase of *Saccharomyces cerevisiae* (Kern et al., 1990). Identical amino acids (*); functionally conserved amino acids (·).

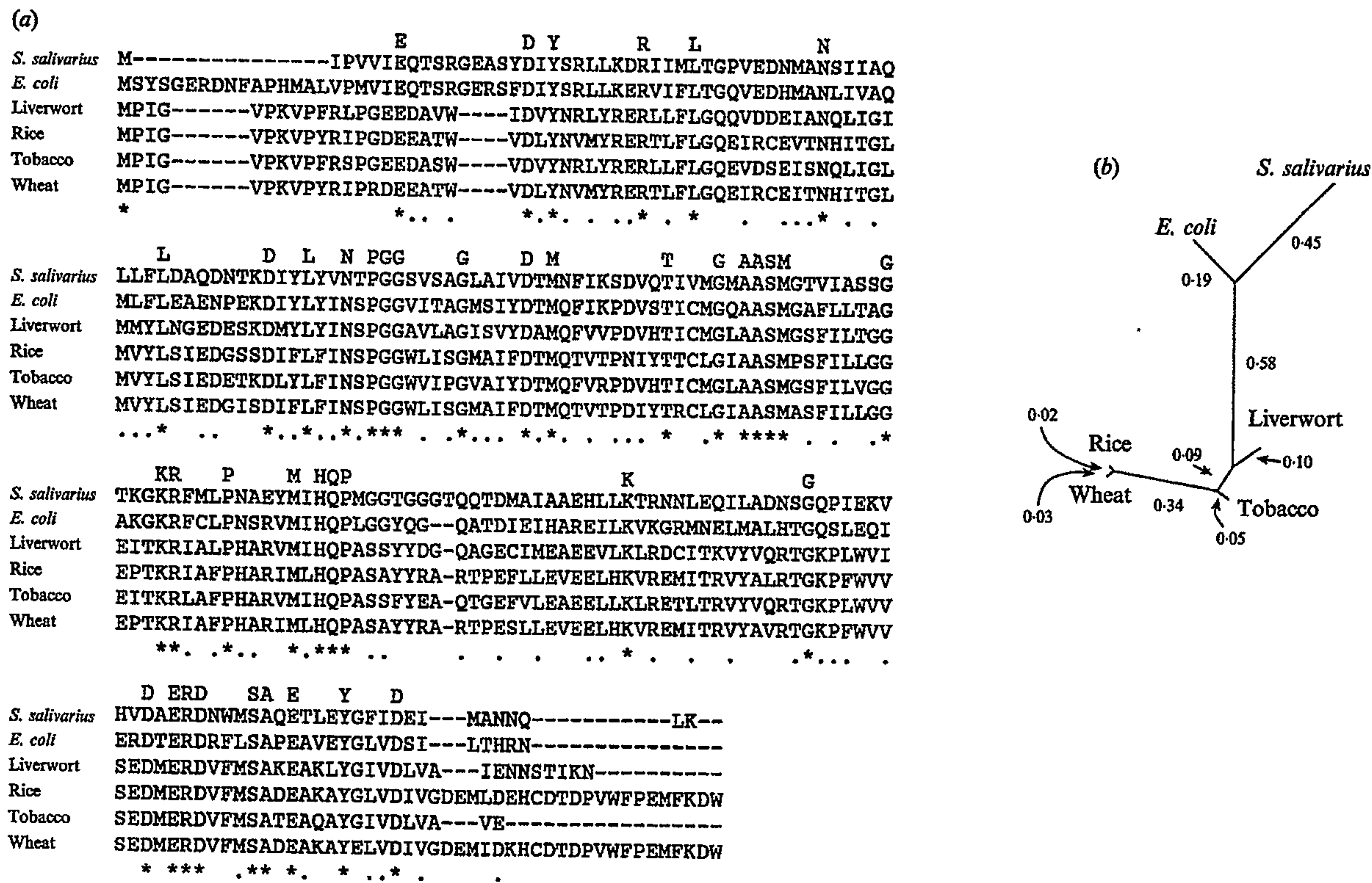


Figure 3.5. (a) Multiple alignment of the deduced amino acid sequence of *orf3* of *S. salivarius* with the ClpPs of *E. coli* (Maurizi *et al.*, 1990a), liverwort (*Marchantia polymorpha*) (Ohyama *et al.*, 1986), rice (*Oryza sativa*) (Hiratsuka *et al.*, 1989), tobacco (*Nicotinia tabacum*) (Shinozaki *et al.*, 1986) and wheat (*Tritichum aestivum*) (Gray *et al.*, 1990); the first methionine being the initiator for each sequence. Sequence conservation is as shown in Fig 3.4, except that identical amino acids are also shown above the six sequences. (b) Unrooted phylogenetic tree of the sequences in (a) constructed using the neighbour-joining method of Saitou and Nei (1987). Relative phylogenetic distances are indicated both numerically and by the length of the connecting lines.

ClpP (Gottesman *et al.*, 1990b; Squires *et al.*, 1991). They are known to act as protein-activated ATPases and are suspected to also encode a chaperonin-like function (Squires and Squires, 1992). The ClpB in both yeast and *E. coli* have been identified as heat shock proteins (Parsell *et al.*, 1991; Kitagawa *et al.*, 1991; Squires *et al.*, 1991). The proximity of both these genes, *clpB* and *clpP* to the *ftf* gene in *S. mutans* and *S. salivarius* respectively, may be significant as there is evidence that when cell-associated Ftf of *S. salivarius* is inactivated by oxygen radicals it is proteolytically removed from the cell surface (Jacques & Wittenberger, 1981; Abbe *et al.*, 1986). The nature of the proteinase(s) involved remains unknown.

The phylogenetic tree originally published following analysis of *orf3* (Giffard *et al.*, 1993; Fig. 3.5b) emphasized that the product of *orf3* could be assigned to the ClpP family. However, it also showed that the *E. coli* and *S. salivarius* sequences were more closely related to each other than to the chloroplast sequences. The topological separation of the bacterial sequences from the chloroplast sequences was tested by the "bootstrap" sampling method (Felsenstein, 1985) using 1000 samples and gave a 100% confidence limit. These results appear to be at variance with current phylogenetic hypotheses since it is generally believed that higher plant chloroplasts arose from cyanobacterial ancestors. According to Woese (1987), there is some evidence that the Gram-positive bacteria and the cyanobacteria form a natural group that does not include the purple bacteria. The purple bacteria include the *Enterobacteriaceae* and species such as *E. coli*. One plausible explanation for this is that the evolutionary pressures on ClpP protease sequences are different in chloroplasts than in free-living bacteria. Alternatively, the chloroplast *clpP* genes may have originated in the plant genome and thus be somewhat phylogenetically remote from the bacterial sequences.

3.3.4 DNA Upstream of the *ftf* Gene Codes for an Insertion Sequence

Considering the association between *clp* homologues and the *ftf* genes in *S. salivarius* and *S. mutans*, it was decided to determine whether the *S. salivarius ftf* was preceded by a *clpB* homologue as it is on the *S. mutans* chromosome.

A stretch of 1101 bp of DNA was sequenced in pKRK108 upstream of the *EcoRI* site in the *S. salivarius* chromosomal insert in λ A-39 (Fig 3.1). Database searches failed to show any homology between this sequenced region and the *clp* gene families. However, the sequenced region did contain the beginning of an ORF that started with an ATG codon and was preceded by a putative ribosome-binding site (*orf4*; Fig. 3.3). Sequence determination was continued 240

bp downstream of the *EcoRI* site and the end of *orf4* was found 46 bp downstream of this *EcoRI* site. This region also contains the putative –10 and –35 promoter sites of the *ftf* gene (Chapter 4; Fig 3.3).

The potential product of *orf4* is 326 amino-acids in length (Fig 3.4). A search of the Genbank translated open reading frame data base initially revealed clear similarity with two transposases - one encoded by the *E. coli* insertion sequence IS30 (Dalrymple *et al.*, 1984) and the other by the clindamycin resistance transposon Tn4551 which is found in the *Bacteroides* plasmid pBI136 (Smith, 1987) (Fig 3.6). Similarity between these two transposable elements had not been previously been noted.

Transposable genetic elements are usually delineated by inverted repeats and flanked by direct repeats. The presence of such repeats is considered diagnostic for a sequence that is the product of a transposition event (Lewin, 1990). Such repeats were found in the expected positions with respect to *orf4*, and these are shown in Fig 3.7. The structure within the boundary of these repeats can therefore be considered to be an insertion sequence-like element and has been named IS1611. One interesting property common to both IS30 and the insertion sequences that delineate Tn4551 is that they can apparently increase the expression of downstream genes (Smith, 1987; Neuwald and Stauffer, 1990). It is not known whether IS1611 also possesses this property, but it is interesting to note that *S. salivarius* ATCC 25975 expresses 3 to 4 times the amount of Ftf as two other strains of *S. salivarius* (Milward and Jacques, 1990).

Subsequent to our reporting of IS1611, an iso-form of this putative insertion element was reported in *S. salivarius* (IS1139; Lortie *et al.*, 1994). Southern hybridization of restriction endonuclease-digested genomic DNA from 21 strains of oral streptococci, using a probe specific to the transposase-encoding gene, *tnpA*, of iso-IS1611 (IS1139) revealed that it was found in only two strains of *S. salivarius*, ATCC 25975 and ATCC 13419, in eight and two copies, respectively (Lortie *et al.*, 1994).



Figure 3.6. Multiple alignment of the deduced amino acid sequence of *orf4* of *S. salivarius* with the presumptive transposases from IS30 (Dalrymple *et al.*, 1984) and Tn4551 (Smith, 1987). Sequence conservation is as shown in Fig 3.4.

CHAPTER 4

THE CELL-ASSOCIATED FRUCTOSYLTRANSFERASE OF *STREPTOCOCCUS SALIVARIUS*¹

4.1 INTRODUCTION

Unlike the bacilli and mutans streptococci which secrete their Ftfs directly into the culture fluid (Carlsson, 1970; Chambert and Petit-Glatron, 1984), the higher molecular weight Ca^{2+} -dependent Ftf of *Streptococcus salivarius* is initially cell-associated (Garszcynski and Edwards, 1973; Jacques and Wittenberger, 1981). This cell-association is lost in the presence of the enzyme's substrate, sucrose, for under these conditions the Ftf is released from the cell and secreted into the culture medium (Milward and Jacques, 1990). As a preliminary to analysing the mode of Ftf cell-attachment and release, the *ftf* of *S. salivarius* ATCC 25975 was manually sequenced. As cell-attachment is usually directed by domains either at the N- or C-terminus of a protein, it was reasoned that alignment of amino acid sequence of the Ftf of *S. salivarius* with those of other bacterial Ftfs should highlight unique regions involved in surface-attachment. Any obvious region of interest would then be targeted for further analysis by sub-cloning or deletion studies. This chapter reports the sequencing of the *ftf* gene and the comparison of the deduced amino acid sequence of the Ftf protein with those of other bacterial Ftfs available at the time.

4.2 METHODS

4.2.1 DNA Sequencing

The *ftf* gene was sequenced using the phagemid, pKRK104, containing the 3.6 kbp *EcoRI*-*Bgl*II fragment of λ A-39 propagated in *E. coli* strain JM109 or NM522 (Table 2.1; Fig 3.1). Sequencing reactions were carried out using CsCl purified dsDNA as described in Sections 2.4.5 and 2.4.16 using Pharmacia T3 and T7 primers as well as custom made oligonucleotide primers where necessary (Table 4.1). The DNA sequence of the insert of *S. salivarius* chromosomal DNA in pKRK104 was confirmed in both directions (Fig 4.1). The nucleotide sequence of the *ftf* gene was submitted to GenBank and assigned the accession numbers L08445.

¹ Some of the results of this chapter have already been published – Rathsam *et al.*, *J. Bacteriol.* 175: 4520-4527, 1993.

Table 4.1 Primers used in the sequencing of the *ftf* gene

Primer*	DNA Sequence
F1 (T3)	TTAACCCTCACTAA
F2	TTCACAGTCAAATACAG
F3	CTCCATCAGTAGTTGAA
F4	GCTATCAAGGGTACAGG
F5	GACTTTACTCATTGGAA
F6	AGACCGTACGGATAACT
F7	GCCACATGGTTACTGAT
F8	GATGAATGCAGATGATA
F9	AGACTAATAAACCAGGC
F10	GCTCCAAAACCAGTTAA
F11	TGGAGTGAGTATAAAGC
F12	CTTTCTGGTAAGAAATT
B1	GCTTCCGGCTCGTATGT
B2 (T7)	AAATTAATACGACTCA
B3	AATTTCTTACCAGAAA
B4	GCTTTATACTCACTCCA
B5	TTAACTGGTTTTGGAGC
B6	GCCTGGTTTATTAGTCT
B7	TATCATCTGCATTCATC
B8	GCAGTGAAGAGGTAATA
B9	AGTTATCCGTACGGTCT
B10	TTCCAATGAGTAAAGTC
B11	CCTGTACCCTTGATAGC
B12	TTCAACTACTGATGGAG
B13	CTGTATTTGACTGTGAA

* (F), Forward primer (5'→3') initiated from the T3 RNA promoter (T3) of pIBI31; (B), Reverse primer (3'→5') initiated from the T7 RNA promoter (T7) of pIBI31

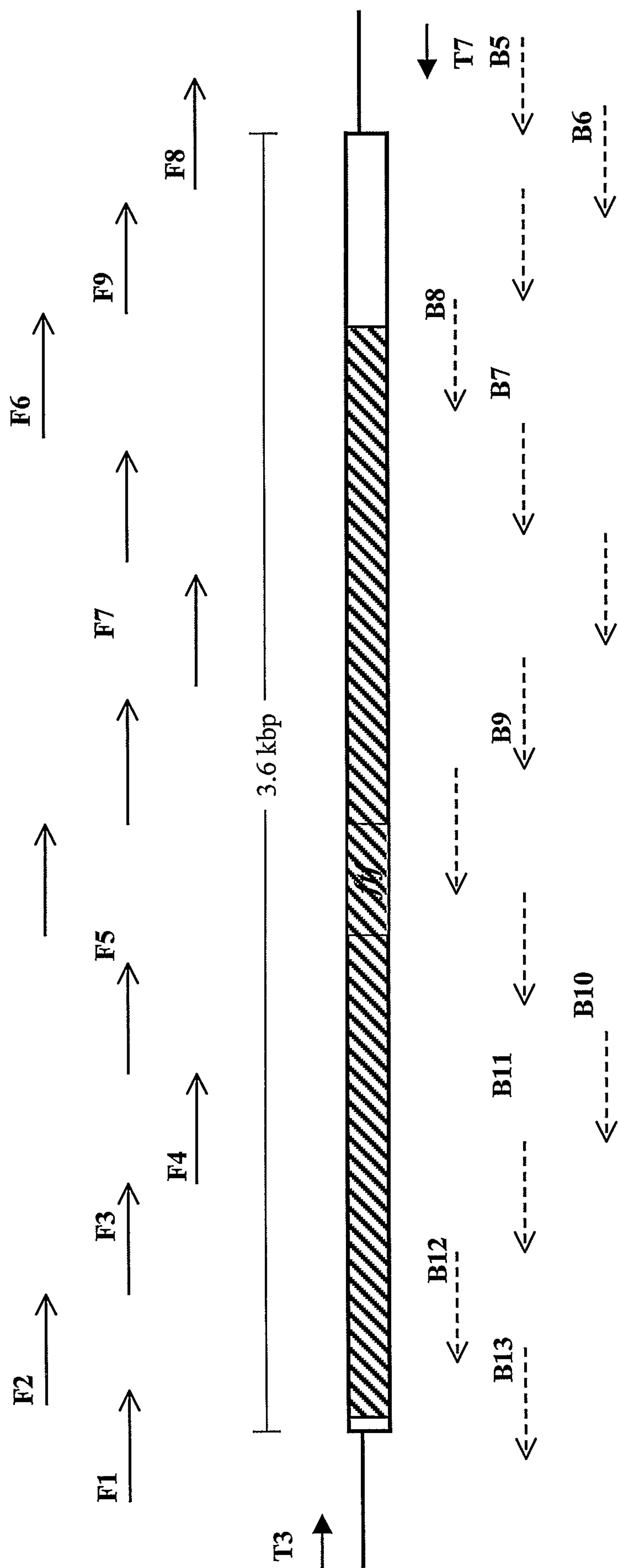


Figure 4.1. Diagram showing the method employed to sequence the 3.6 kbp *EcoRI-BglII* insert containing the *ftf* gene (hatched region); Sanger *et al.*, 1977). Overlapping sequences were obtained by 'gene-walking' at points initiated by synthesized primers shown here by the arrows and described in Table 4.1. The *ftf* gene was found to be transcribed from left to right in the above diagram.

4.2.2 DNA Sequence Analysis

DNA sequences were assembled and ORFs detected using the IBI-Pustell sequence analysis software and database searches and analyses carried out as previously described (2.6).

4.3 RESULTS

4.3.1 Sequencing of the *ftf* Gene of *S. salivarius*

The sequence of the 3,636 bp fragment in pKRR104 was determined in its entirety and found to include an ORF of 2,907 bp beginning with a ATG codon. N-terminal analysis of the secreted Ftf, however, indicated that the mature secreted enzyme devoid of its signal sequence possessed the sequence DQVTE (see Section 7.3.6) indicating that the translation start site, based on the known lengths of streptococcal signal sequences (von Heijne, 1983), occurred at an ATG codon 63 bp downstream from the first of an ATG codon in the ORF. Consequently, the translation start site of the Ftf differs from that originally published (Rathsam *et al.*, 1993; Rathsam and Jacques, 1998), and actually codes for a mature protein with a predicted M_r 98,450. The predicted translation start site was preceded by putative ribosome binding and promoter sequences, and the end of the ORF immediately followed by a potential factor-independent transcription terminator site 8 bp downstream (Fig 4.2).

4.3.2 The *S. salivarius* Ftf Shares Common Ancestry with Equivalent Enzymes

The deduced amino acid sequence of the Ftf of *S. salivarius* was compared with the sequences of other known Ftfs and levansucrases (SacBs) including those from the Gram-positive bacteria, *S. mutans* (Shiroza and Kuramitsu, 1988), *B. subtilis* (Fouet *et al.*, 1984), *B. amiloliquefaciens* (Tang *et al.*, 1990) and *B. stearothermophilus* (Li *et al.*, 1997), as well as the Gram-negative bacteria, *E. amylovora* (Geier and Geider, 1993), *Z. mobilis* (Song *et al.*, 1993) and *A. diazotrophicus* (Arrieta *et al.*, 1996) (Fig 4.3). Several important amino acids previously identified in the SacB of *B. subtilis* were identical in the Ftf of the Gram-positive bacterium, *S. salivarius*, and/or conserved in the Gram-negative bacteria (Figs 4.2 & 4.3). These included R⁵¹³ (*S. salivarius* numbering) which is considered to act as the proton donor in fructan synthesis (Chambert and Petit-Glatron, 1991) and also G⁵⁵⁵ which produces a β -turn apparently required for structural integrity of the enzymes as well as for the release of the SacB from the cytoplasmic membrane of *B. subtilis* during the secretion process (Petit-Glatron *et al.*, 1990). Interestingly, the suggestion by Chambert and Petit-Glatron (1991) that the synthesis of inulin by the Ftf of *S. mutans* is reflected in an alteration in the amino acid sequence on either side of R⁵¹³ would not appear to be the case, as this region

Figure 4.2 Sequence of the DNA insert in pKRK104 in the 5'→3' direction containing the *ftf* gene and its translated open reading frame. The numbers refer to the base pair of the chromosomal insert of *S. salivarius* DNA originally cloned into λA-39 (Table 2.1; Chapter 3). A putative ribosome-binding site (RBS; blue colour), putative promoter site.(–10 /-35, red colour), the signal sequence cleavage site (P1), and the pro-peptide cleavage site (P2) are shown.. The sites within the *ftf* gene where 3'-deletions were made are indicated by the name of the plasmid containing the deletion, while the beginnings and ends of the wall-spanning and membrane spanning regions are also indicated. The underlined region in bold indicates the site of a potential down-stream factor independent transcriptional terminator site.

1250	1260	1270	-35	1290	1300	-10
A	AGATTGGAAT	TTAGATAAAA	AAATGATAGA	ATGGAACATG	TAAAAACCTA	AAAGGAGAGT
1320	1330	1340	1350	1360	RBS	
TTATATGGAT	ATTACAGTTA	ATTCACAGTC	AAATACAGTA	GCACCAAAGC	AAGCAGAGTG	
1380	1390	1400	1410	1420		
CAAAAAG	ATG CGT TAT	AGC ATC CGT	AAG GTA GCC	ACA GTA GGG	GCA ACA TCA	
	Met Arg Tyr	Ser Ile Arg	Lys Val Ala	Thr Val Gly	Ala Thr Ser	
1430	1440	1450	1460	1470	↓ P1	
GCA CTC GTT	GGT ACC TTG	GCG TTT TTG	GGT GCC ACT	CAG GTA AAA	GCT GAT	
Ala Leu Val	Gly Thr Leu	Ala Phe Leu	Gly Ala Thr	Gln Val Lys	Ala Asp	
1480	1490	1500	1510	1520		
CAA GTT ACA	GAA ACA GCA	CCA GCG GTA	GCT ACT GCG	ACA GCA ACA	CCA GAA	
Gln Val Thr	Glu Thr Ala	Pro Ala Val	Ala Thr Ala	Thr Ala Thr	Pro Glu	
1530	1540	1550	1560	1570		
ACA AGC ACA	GCA TCT CTA	ACG GTA GCC	AGT GAA ACA	GCA ACA AGT	GTG GCA	
Thr Ser Thr	Ala Ser Leu	Thr Val Ala	Ser Glu Thr	Ala Thr Ser	Val Ala	
1580	1590	1600	1610	1620		
ACT TCA GAA	GCA GTT GAA	TCT TCA GTT	GCA CAT TCA	GAG GTT GCA	ACT AAA	
Thr Ser Glu	Ala Val Glu	Ser Ser Val	Ala His Ser	Glu Val Ala	Thr Lys	
1630	1640	1650	1660	1670		
CCA GTT ACA	GAG ACA CAA	CCA AGC AAC	ACA ACT CCA	TCA GTA GTT	GAA GAA	
Pro Val Thr	Glu Thr Gln	Pro Ser Asn	Thr Thr Pro	Ser Val Val	Glu Glu	
1680	1690	1700	↓ P2	1710	1720	
AAG GCG TCT	AGC ACA GTA	GTG ACA TCT	TCA TCA GAT	GCA ACA ACA	CCT TCA	
Lys Ala Ser	Ser Thr Val	Val Thr Ser	Ser Ser Ser	Asp Ala Thr	Thr Pro Ser	
1730	1740	1750	1760	1770		
GCG ACT GTT	GCT GCA GTT	TCT GCT CCG	GCT CAT ACA	TCA GAG GCT	GCA GTA	
Ala Thr Val	Ala Ala Val	Ser Ala Pro	Ala His Thr	Ser Glu Ala	Ala Val	
1780	1790	1800	1810	1820	1830	
GAA GCA CCA	ACT TCA ACA	GCA AGC TCA	GAA GCA GCT	GAT ACA CAC	ACT GAA	
Glu Ala Pro	Thr Ser Thr	Ala Ser Ser	Glu Ala Ala	Asp Thr His	Thr Glu	
1840	1850	1860	1870	1880		
GTA GAC CTT	AAA GTT TCT	GAA AAC AGC	GCG GCT AAT	GCC AAC CTT	AGT AAA	
Val Asp Leu	Lys Val Ser	Glu Asn Ser	Ala Ala Asn	Ala Asn Leu	Ser Lys	
1890	1900	1910	1920	1930		
CTT AAT GGT	CGT ATC AAG	TCT ATT GTT	GAA GAG AAT	ATG ACC TCA	GAT CAA	
Leu Asn Gly	Arg Ile Lys	Ser Ile Val	Glu Glu Asn	Met Thr Ser	Asp Gln	
1940	1950	1960	1970	1980		
ATC GTT GCC	TTG ACT GAG	GAA GAA ATT	AAA GCA CTT	AAT AAA GTT	GAC TTT	
Ile Val Ala	Leu Thr Glu	Glu Glu Ile	Lys Ala Leu	Asn Lys Val	Asp Phe	

Figure 4.2 (cont)

	1990			2000			2010			2020			2030				
	AGT	GAT	GAT	GCT	ATC	AAG	GGT	ACA	GGT	ACT	AGC	TTG	ACT	TAC	CGT	AAC	TTG
	Ser	Asp	Asp	Ala	Ile	Lys	Gly	Thr	Gly	Thr	Ser	Leu	Thr	Tyr	Arg	Asn	Leu
	2040			2050			2060			2070			2080				
	AAA	GAT	ATC	GTT	GCT	AGC	TTC	TTG	AAG	CAA	GAC	AGC	AAA	TTG	GCA	GTT	CCT
	Lys	Asp	Ile	Val	Ala	Ser	Phe	Leu	Lys	Gln	Asp	Ser	Lys	Leu	Ala	Val	Pro
	2090			2100			2110			2120			2130				
	TAC	TTT	AAA	GCA	GAT	ACA	ATC	ATC	AAT	ATG	CCT	GCC	TTC	AAC	ACT	GTT	GAT
	Tyr	Phe	Lys	Ala	Asp	Thr	Ile	Ile	Asn	Met	Pro	Ala	Phe	Asn	Thr	Val	Asp
	2140			2150			2160			2170			2180				
	GCC	CAA	ACG	ATG	AAA	AAA	GAA	GAA	ATC	GAC	GTT	TGG	GAT	TCT	TGG	CCA	GTC
	Ala	Gln	Thr	Met	Lys	Lys	Glu	Glu	Ile	Asp	Val	Trp	Asp	Ser	Trp	Pro	Val
2190	2200			2210			2220			2230							
	CAA	GAT	GCT	AAG	TCA	GGT	GTA	GTT	AGC	AAC	TGG	AAT	GGT	TAC	CAG	TTG	GTT
	Gln	Asp	Ala	Lys	Ser	Gly	Val	Val	Ser	Asn	Trp	Asn	Gly	Tyr	Gln	Leu	Val
2240	2250			2260			2270			2280							
	ATC	TCA	ATG	GCC	GGT	GCA	CCA	AAT	AAA	AAT	TCA	AAC	CAC	ATT	TAC	TTG	CTC
	Ile	Ser	Met	Ala	Gly	Ala	Pro	Asn	Lys	Asn	Ser	Asn	His	Ile	Tyr	Leu	Leu
2290	2300			2310			2320			2330			2340				
	TAT	AGA	AAA	TAT	GGG	GAT	AAT	GAC	TTT	ACT	CAT	TGG	AAA	AAT	GCG	GGT	CCA
	Tyr	Arg	Lys	Tyr	Gly	Asp	Asn	Asp	Phe	Thr	His	Trp	Lys	Asn	Ala	Gly	Pro
	2350			2360			2370			2380			↓ G ³⁰⁵ 2390				
	ATC	TTT	GGT	TAC	AAT	GCC	CTT	GAA	GAT	GAC	CAA	CAA	TGG	TCA	GGT	TCA	GCG
	Ile	Phe	Gly	Tyr	Asn	Ala	Leu	Glu	Asp	Asp	Gln	Gln	Trp	Ser	Gly	Ser	Ala
	2400			↓ D ³¹²			2420			2430			2440				
	ACT	GTT	AAC	TCT	GAT	GGT	AGT	ATC	CAA	CTT	TAT	TAC	ACT	AAG	AAT	GAT	ACT
	Thr	Val	Asn	Ser	Asp	Gly	Ser	Ile	Gln	Leu	Tyr	Tyr	Thr	Lys	Asn	Asp	Thr
	2450			2460			2470			2480			2490				
	AGT	GGT	GGT	AAA	TTA	AAC	TGG	CAA	CAA	TTG	GCT	AGT	GCT	ACA	CTT	AAC	TTG
	Ser	Gly	Gly	Lys	Leu	Asn	Trp	Gln	Gln	Leu	Ala	Ser	Ala	Thr	Leu	Asn	Leu
	2500			2510			2520			2530			2540				
	GCT	GTT	GAA	AAC	GAC	GAA	GTT	GTT	ATC	AAG	TCA	GTT	GAA	AAT	GAC	CAC	ATT
	Ala	Val	Glu	Asn	Asp	Glu	Val	Val	Ile	Lys	Ser	Val	Glu	Asn	Asp	His	Ile
	2550			2560			2570			2580			2590				
	CTC	TTT	GGT	GGA	GAT	AAT	TAT	CAC	TAT	CAA	AGC	TAT	CCT	AAA	TTT	ATG	AGT
	Leu	Phe	Gly	Gly	Asp	Asn	Tyr	His	Tyr	Gln	Ser	Tyr	Pro	Lys	Phe	Met	Ser
	2600			2610			2620			2630			2640				
	ACT	TTT	GAT	GAT	GAC	CAT	AAC	CAT	GAT	GGT	AAT	CCA	GAC	CGT	ACG	GAT	AAC
	Thr	Phe	Asp	Asp	Asp	His	Asn	His	Asp	Gly	Asn	Pro	Asp	Arg	Thr	Asp	Asn
	2650			↓ D ³⁹⁷			2670			2680			2690				
	TAC	TGT	CTT	CGT	GAC	CCA	CAC	ATC	ATT	GAA	GAT	AAC	GGT	AGC	CGT	TAC	CTT
	Tyr	Cys	Leu	Arg	Asp	Pro	His	Ile	Ile	Glu	Asp	Asn	Gly	Ser	Arg	Tyr	Leu
2700	2710			2720			2730			2740							
	ATT	TTT	GAA	TCA	AAT	ACA	GGT	GAC	GAA	AAC	TAC	CAA	GGC	GAA	AAA	CAA	ATT
	Ile	Phe	Glu	Ser	Asn	Thr	Gly	Asp	Glu	Asn	Tyr	Gln	Gly	Glu	Lys	Gln	Ile
2750	2760			2770			2780			2790							
	TAC	AAA	TGG	TCA	AAC	TAT	GGA	GGA	GAT	GAC	GCC	TTT	AAC	CTT	AAG	TCA	TTC
	Tyr	Lys	Trp	Ser	Asn	Tyr	Gly	Gly	Asp	Asp	Ala	Phe	Asn	Leu	Lys	Ser	Phe
2800	2810			2820			2830			2840			2850				
	CTC	AAT	ATT	GTT	AAT	AAT	AAA	CAC	CTT	TAC	AAC	CTT	GCA	TCA	TGG	GCA	AAT
	Leu	Asn	Ile	Val	Asn	Asn	Lys	His	Leu	Tyr	Asn	Leu	Ala	Ser	Trp	Ala	Asn
	2860			2870			2880			2890			2900				
	GGT	TCT	ATC	GGT	ATC	TTG	AAA	CTT	GAT	GAT	AAT	GAG	AAA	AAT	CCA	TCT	GTA
	Gly	Ser	Ile	Gly	Ile	Leu	Lys	Leu	Asp	Asp	Asn	Glu	Lys	Asn	Pro	Ser	Val

Figure 4.2 (cont)

2910	2920	2930	2940	2950	
GCT GAA CTT TAC ACA CCG CTT GTC ACT AGC CAC ATG GTT ACT GAT GAA GTA					
Ala Glu Leu Tyr Thr Pro Leu Val Thr Ser His Met Val Thr Asp Glu Val					
2960	2970	2980	2990	3000	
GAA CGT CCA AGT GTT GTT AAG ATG GGT GGC AAG TAT TAC CTC TTC ACT GCT					
Glu Arg Pro Ser Val Val Lys Met Gly Gly Lys Tyr Tyr Leu Phe Thr Ala					
↓ R ⁵¹³	3020	3030	3040	3050	
TCA CGT ATC AAC AAA TCT ACG GAT GCT GAA GGT ACT GTA GCA GCG CGT GAA					
Ser Arg Ile Asn Lys Ser Thr Asp Ala Glu Gly Thr Val Ala Ala Arg Glu					
3060	3070	3080	3090	3100	
GCT GTC GGT GAT GAC GTT GTC ATG CTT GGA TTC GTA TCA GAC AGC CTC CGA					
Ala Val Gly Asp Asp Val Val Met Leu Gly Phe Val Ser Asp Ser Leu Arg					
3110	3120	3130	↓ G ⁵⁵⁵	3150	
GGA AAA TAT CGT CCA CTT AAT GGC TCA GGT GTG GTC TTG ACA GCC TCT GTA					
Gly Lys Tyr Arg Pro Leu Asn Gly Ser Gly Val Val Leu Thr Ala Ser Val					
3160	3170	3180	3190	3200	
CCA GCT GAC TGG CGT ACA TCT ACC TAT TCT TAC TAT GCG GTT CCT GTT GAA					
Pro Ala Asp Trp Arg Thr Ser Thr Tyr Ser Tyr Tyr Ala Val Pro Val Glu					
3210	3220	3230	3240	3250	
GGC TCA TCA GAT ACT CTT TTG GTA ACT TCT TAC ATG ACT AAC CGT GGT GGA					
Gly Ser Ser Asp Thr Leu Leu Val Thr Ser Tyr Met Thr Asn Arg Gly Gly					
3260	3270	3280	3290	↓ Deletion in pKRK1661(HindIII)	
ATT GCC GGT GCT GAG AAT AAA TCA ACA TGG GCT CCA AGC TTC CTT ATC AAG					
Ile Ala Gly Ala Glu Asn Lys Ser Thr Trp Ala Pro Ser Phe Leu Ile Lys					
3310	3320	3330	3340	3350	3360
ATG AAT GCA GAT GAT ACT ACA GAA GTT CTT CCT AAG ATG ACG AAT CAA GGT					
Met Asn Ala Asp Asp Thr Thr Glu Val Leu Pro Lys Met Thr Asn Gln Gly>					
3370	3380	3390	3400	3410	
GAC TGG ATT TGG GAT AAA TCT AGT GAA AGC TTG GTA CAC GTA GGC GAT CAA					
Asp Trp Ile Trp Asp Lys Ser Ser Glu Ser Leu Val His Val Gly Asp Gln					
3420	3430	3440	3450	3460	
AAC AGT GCT AAA TTG CCA AAC GAG GAC TTT AAT GTA GAT TAT TAT GCG GTA					
Asn Ser Ala Lys Leu Pro Asn Glu Asp Phe Asn Val Asp Tyr Tyr Ala Val					
3470	3480	3490	3500	3510	
TCG GGT TAT GGT CTT AAA CCA CAT ACT TAT CCT ACA GTT GAT GGT TCA ACA					
Ser Gly Tyr Gly Leu Lys Pro His Thr Tyr Pro Thr Val Asp Gly Ser Thr					
3520	3530	3540	3550	3560	
GGC GTT TCT GAA GCT CAC GGT GTC TTG ACT GTA ACT GTT AAA GAT GGT AAA					
Gly Val Ser Glu Ala His Gly Val Leu Thr Val Thr Val Lys Asp Gly Lys					
3570	3580	↓ Wall Domain Commences	3610		
GAC AAG AAG GCA GAT AAG CCA GAA ACA CCA GTA AGT CCA ACA GAA GGT AAC					
Asp Lys Lys Ala Asp Lys Pro Glu Thr Pro Val Ser Pro Thr Glu Gly Asn					
3620	3630	3640	3650	3660	
CAT TCT GTT GAT GAT AAG ACT AAT AAA CCA GGC ACT TCT AAA CCA GCG GAC					
His Ser Val Asp Asp Lys Thr Asn Lys Pro Gly Thr Ser Lys Pro Ala Asp					
3670	3680	3690	3700	3710	
AAT AAT CAA CCA TCA GCT GAT AAA GAA GAT AAA CCA ACA AAT CCT ACG AAT					
Asn Asn Gln Pro Ser Ala Asp Lys Glu Asp Lys Pro Thr Asn Pro Thr Asn					
3720	3730	3740	3750	↓ Deletion in pKRK1816	
CCG GAT TCT CCA GCT CGT ACT CCT TTC CCA TAT TAT GGA GAC CAC TCT AAT					
Pro Asp Ser Pro Ala Arg Thr Pro Phe Pro Tyr Tyr Gly Asp His Ser Asn					
3770	3780	3790	3800	3810	
GAT AAC AAT TCT TCA AAT GAT CAT CAT GTA GCG GTC CCT GTG AAA CCT TCG					
Asp Asn Asn Ser Ser Asn Asp His His Val Ala Val Pro Val Lys Pro Ser					

Figure 4.3 Multiple alignment of the deduced amino acid sequence of the Ftf of *S. salivarius* (pFtf_Ssal; Accession No: L08445) with that of the Ftfs of *S. mutans* (pFtf_Smut; Accession No: P11701), *B. subtilis* (pFtf_Bsub; Accession No: A25040), *B. amyloliquifaciens* (pFtf_Bamy; Accession No: JQ0802), *B. stearothermophilus* (pFtf_Bste; Accession No: AAB97111), *E. amylovora* (pFtf_Eamy; S39195), *Z. mobilis* (pFtf_Zmob; Accession No: AAA27695) and *A. diazotrophicus* (pFtf_Adia; Accession No: Q43998). Key amino acids are indicated in bold. Conserved amino acids in the “sucrose box” are shaded. Identical amino acids (*); functionally conserved amino acids (•).

pFtf_Ssal	M R Y S I R K V A T V G A T S A L V G T L A F L G A T Q V K A D Q V T E T A P A V A T A T A T P E T S T A S L T V A S E
pFtf_Smut	M E T K V R K K M Y K ----- K G K F W V V A T I T T A M L T G I G L S S V Q A D E A N S T Q V S S E L
pFtf_Bsub	-- M N I K K F A K Q ----- A T V L T F T T A L L A G G A T Q A F A K E T N -----
pFtf_Bamy	-- M N I K K I V K Q ----- A T V L T F T T A L L A G G A T Q A F A K E N N -----
pFtf_Bste	-- M N I K K F A K Q ----- A T V L T F T T A L L A G G A T Q A F A K E T N -----
pFtf_Eamy	----- M S D Y N -----
pFtf_Zmob	----- M L N K A G I -----
pFtf_Adia	M A H V R R K V A T L N ----- M A L A G S L L M V L G A Q S A L A Q G N F S R Q E A A R M A H R P G V M P R G
pFtf_Ssal	T A T S V A T S E A V E S S V A H S E V A T K P V T E T Q P S N T T P S V V E E K A S S T V V T S S S D A T T P S A T V
pFtf_Smut	A E R S Q V Q E N T T A S S S A A E N Q A K T E V Q E T P S T N P A A A T V E N T D Q T T K V I T D N A A V E S K A S K
pFtf_Bsub	----- Q K P Y K E T Y G -----
pFtf_Bamy	----- Q K A Y K E T Y G -----
pFtf_Bste	----- Q K P Y K E T Y G -----
pFtf_Eamy	----- Y K P T L W T R A -----
pFtf_Zmob	----- A E P S L W T R A -----
pFtf_Adia	G P L F P G R ----- S L A G V P G F P L P S I H T Q Q A Y D P Q S D -----
pFtf_Ssal	A A V S A P A H T S E A A V E A P T S T A S S - E A A D T H T E V D L K V S E N S A A N A N L S K L N G R I K S I V E E
pFtf_Smut	T K D Q A A T V T K T A A S T P E V G Q T N E K D K A K A T K E A D I T T P K N T I D E Y G L T E Q A R K I A T E A G I
pFtf_Bsub	----- I S H I T R H D M L Q I P E Q -----
pFtf_Bamy	----- V S H I T R H D M L Q I P K Q -----
pFtf_Bste	----- I S H I T R H D M L Q I P E Q -----
pFtf_Eamy	----- D A L K V H E D -----
pFtf_Zmob	----- D A M K V H T D -----
pFtf_Adia	----- F T A R W T R A D A L Q I K A H S D A T V A A G Q N -----
pFtf_Ssal	N M T S D Q I V A L T E E E I K A L N K V D F S D D A I K G T G T S L T Y R N L K D I V A S F L K Q D S K L A V P Y F K
pFtf_Smut	N L S S ----- L T Q K Q V E A L N K V K L T S D A Q -- T G H Q M T Y Q E F D K I A Q T L I A Q D E R Y A I P Y F N
pFtf_Bsub	----- Q K N E K Y Q V P E F D -----
pFtf_Bamy	----- Q Q N E K Y Q V P Q F D -----
pFtf_Bste	----- Q K N E K Y Q V P E F D -----
pFtf_Eamy	----- D P T T T Q P V I D -----
pFtf_Zmob	----- D P T A T M P T I D -----
pFtf_Adia	----- S L P A Q L T M P N I P -----
pFtf_Ssal	A D T I I N M P A F N T V D A Q T M K K E E I D V W D S W P V Q D A K S G V V S N W N G Y Q L V I S M A G A P - N K N -
pFtf_Smut	A K A I K N M K A A T T R D A Q T G Q I A D L D V W D S W P V Q D A K T G E V I N W N G Y Q L V V A M M G I P - N T N -
pFtf_Bsub	S S T I K N I S S A K G ----- L D V W D S W P L Q N A D - G T V A N Y H G Y H I V F A L A G D P K N A D -
pFtf_Bamy	Q S T I K N I E S A K G ----- L D V W D S W P L Q N A D - G T V A E Y N G Y H V V F A L A G S P K D A D -
pFtf_Bste	S S T I K N I S S A K G ----- L D V W D S W P L Q N A D - G T V A N Y H G Y H I V F A L A G D P K N A D -
pFtf_Eamy	I A F P V M S E E V F I ----- W D T M P L R D F D - G E I I S V N G W C I I F T L T A D R N T D N P
pFtf_Zmob	Y D F P V M T D K Y W V ----- W D T W P L R D I N - G Q V V S F Q G W S V I F A L V A D R T K Y G -
pFtf_Adia	A D F P V I N P D V W V ----- W D T W T L I D K H - A D Q F S Y N G W E V I F C L T A D P N A G Y G
pFtf_Ssal	----- S N H I Y L L Y R K Y G D N D F T H ----- W K N A G P I F G ----- Y N A L
pFtf_Smut	----- D N H I Y L L Y N K Y G D N N F D H ----- W K N A G S I F G ----- Y N E T
pFtf_Bsub	----- D T S I Y M F Y Q K V G E T S I D S ----- W K N A G R V F K D S D K F D A N
pFtf_Bamy	----- D T S I Y M F Y Q K V G D N S I D S ----- W K N A G R V F K D S D K F D A N
pFtf_Bste	----- D T S I Y M F Y Q K V G E T S I D S ----- W K T P G R V F K D S D K F D A N
pFtf_Eamy	Q F Q D E N G N Y D I T R D W E D R H G R A R I C Y W Y S R T G K D ----- W I F G G R V M A ----- E G V A
pFtf_Zmob	----- W H N R N D G A R I G Y F Y S R G G S N ----- W I F G G H L L K ----- D G A N
pFtf_Adia	----- F D D R H V H A R I G F F Y R R A G I P A S R R P V N G G W T Y G G H L F P D G A S A Q V Y

Figure 4.3 (cont)

	G³⁰⁵ D³¹²	
pFtf_Ssal	E-----DDQQWSGSATVNSD-GS-IQLYYTKNDTSGGKLNW-----QQLASATLNLAVE	
pFtf_Smut	P-----LTQEWSGSATVNED-GS-LQLFYTKVDTSDKNSNN-----QRLATATVNLGFD	
pFtf_Bsub	DSILKDQTQEWSGSATFTSD-GK-IRLFYT--DFSGKHYGK-----QTLTTAQVNVSAS	
pFtf_Bamy	DPILKDQTQEWSGSATFTSD-GK-IRLFYT--DYSKGHYGK-----QSLTTAQVNVSAS	
pFtf_Bste	DSILKDQTQEWSGSATFTSD-GK-IRLFYT--DFSGKHYGK-----QTLTTAQVNVSAS	
pFtf_Eamy	P-----TTREWAGTPILLNDRGD-IDLYYT--CVTPG-----ATIAKVRGKIVTS	
pFtf_Zmob	P-----RSWEWSGCTIMAPGTANSVEVFFT--SVNDTPSE-----SVPAQCKGYIYAD	
pFtf_Adia	AGQITYTNQAEWSGSSRLMQIHGNTVSVFYTDVAFNRDANANNITPPQAIITQTLGRIHAD	
	. * . * *	
		D³⁹⁷
pFtf_Ssal	NDEVVIKSVENDHILFGG--DNYHYQSYPKFMSTFDDHNDGNPDRTDNYCLRDPHIIE	
pFtf_Smut	DQDVRIILSVENDKVLTPEGVMAYHYQSYQQWRSTFTG-----ADNIAMRDPHVIE	
pFtf_Bsub	DSSLNINGVEDYKSIFDG--DGKTYQNVQQFIDEGNYSS-----GDNHTLRDPHYVE	
pFtf_Bamy	DDTLKINGVEDHKTIFDG--DGKTYQNVQQFIDEGNYTS-----GDNHTLRDPHYVE	
pFtf_Bste	DSSLNINGVEDYKSIFDG--DSKTYQNVQQFIDEGNYSS-----GDNHTLRDPHYVE	
pFtf_Eamy	DQSVSLEGFQQVTSLSFA--DGTIYQTEEQNA---FWN-----FRDPSPTFI	
pFtf_Zmob	DKSVWFDGFDKVTDLFQA--DGLYYADYAENN---FWD-----FRDPHFVI	
pFtf_Adia	FNHVWFTGFTAHTPLLQP--DGVLYQNGAQNE---FFN-----FRDPFTFE	
	. * . * *	
pFtf_Ssal	D---NGSRYLIFESNTGDEN-YQGEKQIYKWSNYGGD-DAFNLKSFLNIWNKHLNHLAS	
pFtf_Smut	DE---NGDRYLVEFASTGTEN-YQGEDQIYNFTNYGGS-SAYNVKSLFRFLDDQDMYNRAS	
pFtf_Bsub	D---KGHKYLVFEANTGTEDGYQGEESLFNKAYYGKS-TSFFRQESQKLLQS-DKKRTAE	
pFtf_Bamy	D---KGHKYLVFEANTGTENGYYQGEESLFNKAYYGGG-TNFFRKESQKLLQS-AKKRDAE	
pFtf_Bste	D---KGHKYLVFEANTGTEDGYQGEESLFNKAYYGKS-TSFFRQESQKLLQS-DKNRTAE	
pFtf_Eamy	DR-NDGKLYMLFEGNVAGPR---GSHEIT-QAEMG-----NVPPGYEDVGG-----AK	
pFtf_Zmob	NP-EDGKTYALFEGNVAMER---GTVAVG-EEEIG-----PVPPKTETPDG-----AR	
pFtf_Adia	DPKHGPNVNYMVEGNTAGQR---GVANCT-EADLGFRPNPNAETLQEVLDG-----GAY	
	* * . * . * . . *	
		R⁵¹³
pFtf_Ssal	WANGSIGILKLDDNEKNPSVAELYTPLVTSHMVTDEVERPSVVKMGKYYLFTASRINKS	
pFtf_Smut	WANAAIGILKLKGDKKTPEVDQFYTPLSSSTMVSELERPNVVKLGDKYYLFTASRLNHG	
pFtf_Bsub	LANGALGMIELNDDY---TLKKVMKPLIASNTVTDEIERANVFKMNGKWYLFSTDSRGSK-	
pFtf_Bamy	LANGALGMIELNDDY---TLKKVMKPLITSNTVTDEIERANVFKMNGKWYLFSTDSRGSK-	
pFtf_Bste	LANGALGMIELNDDY---TLKKVMKPLIASNTVTDEIERANVFKMNGKWYLFSTDSRGSQ-	
pFtf_Eamy	YQAGCVGLAVAKDLG--SEWQILPPLITAVGVNDQTERPHFVFQDGKYYLFTISHKYT-	
pFtf_Zmob	YCAAAGIAQALNEAR--TEWKLLPPLVTAFGVNDQTERPHVVFQNGLYLFTISHHST-	
pFtf_Adia	YQKANIGLAIATDST--LSKWKFLSPLISANCVNDQTERPQVYLHNGKYYIFTISHRTT-	
	. * . * *	
		G⁵⁵⁵
pFtf_Ssal	TDAEGTVAAREAVGDDVVMLGFVS-DSLRGKYRPLN-GSGVVLTAASV-----	
pFtf_Smut	SNNDAWNKANEEVVGDNVVMLGYVS-DQLINGYKPLN-NSGVVLTAASV-----	
pFtf_Bsub	-----MTIDGITSNDIYMLGYVS-NSLTGPYKPLN-KTGLVLKMDL-----	
pFtf_Bamy	-----MTIDGINSNDIYMLGYVS-NSLTGPYKPLN-KTGLVLQMGH-----	
pFtf_Bste	-----MTIDGITSNDIYMLGYVS-NSLTGPYKPLN-KTGLVLKMDL-----	
pFtf_Eamy	-----FADNLTGPD-GVYGFVS-DKLTGPYTPMN-SSGLVLGN-----	
pFtf_Zmob	-----YADGLSGPD-GVYGFVSENGIFGPYEPLN-GSGLVLGN-----	
pFtf_Adia	-----FAAGVDGPD-GVYGFVG-DGIRSDFQPMNYGSLTMGNPTDLNTAAGTDFDPS	
	. * . * *	
pFtf_Ssal	--PADWRTSTYSYAVPVEGSSDTLLVTSYMINRGGIAGAENKSTWAPSFLIKMNADDTT	
pFtf_Smut	--PADWRTATYSYAVPVAGSSDTLLMTAYMINRNEVAGKGNSTWAPSFLIQVLPDGT	
pFtf_Bsub	--DPNDVTFTYSHFAVPQAKG-NNVVITSYMINRGGFYA--DKQSTFAPSFLNIKGKKT	
pFtf_Bamy	--DPNDVTFTYSHFAVPQAKG-NNVVITSYMINRGGFE--DKKATFGPSFLMNIKGNKTS	
pFtf_Bste	--DPNDVTFTYSHFAVPQATG-NNVVITSYMINRGGFYA--DKQSTFAPSFLNIQGGKTS	
pFtf_Eamy	--PSSQPFQTYSHYVMPNGLV-TSFIDSVPWKG-KDYR---IGGTEAPTIVKILLKGRS-	
pFtf_Zmob	--PSSQPYQAYSHYVMINGLV-TSFIDTIPSSDPNVYR---YGGTLAPTIVKLELVGHR-	
pFtf_Adia	PDQNPRAFQSYSHYVMPGGLV-ESFIDTVENRR-----GGTLAPTIVRVRIAQNAS-	
	. * . * *	
pFtf_Ssal	EVLPMKMNQGDWIWDKSSSESLVHVGDNQNSAKLPNE-DFNVDDYYAVSGYGLKPHTYPTVDG	
pFtf_Smut	KVLAEMTQQGDWIWDEPSRTTDTVGTLDTAYLPGENDGYIDWNVIGGYGLKPHTPGQYQP	
pFtf_Bsub	VVKDSILEQG-----QLTVNK-----	
pFtf_Bamy	VVKNSILEQG-----QLTVN-----	
pFtf_Bste	VVKASILDQG-----QLTVNQ-----	
pFtf_Eamy	-FIVDSFDYG-----YIPAMKDITLK-----	
pFtf_Zmob	-FVTEVKGYG-----YIPPQIEWLAEDSSNSAAALSLLNK-----	
pFtf_Adia	AVDLRYGNNG-----LGGYGDIPANRADVNIAGFIQDLFG-----QPTSGLAQA	
	. * . * *	

Figure 4.3 (cont)

pFtf_Ssal	STGVSEAHGVLTVTVKDGKDKKADKPETPVSPTEGNHSVDDKTNKPGTSKPADNNQPSAD
pFtf_Smut	TVPSTPIHTDDIISFEVSFDG-----HLVIKPVKVNND SAG
pFtf_Bsub	-----
pFtf_Bamy	-----
pFtf_Bste	-----
pFtf_Eamy	-----
pFtf_Zmob	-----
pFtf_Adia	STNNAQVLAQVRQFLNQ-----
pFtf_Ssal	KEDKPTNPTNPDSPARTPFPYYGDHSNDNNSNDHHVAVPVK PSTGDSVGDRRPVAQAAE
pFtf_Smut	RIDQ-----S-----RNSGGS LNVA FNVSAGGNISVK-----
pFtf_Bsub	-----
pFtf_Bamy	-----
pFtf_Bste	-----
pFtf_Eamy	-----
pFtf_Zmob	-----
pFtf_Adia	-----
pFtf_Ssal	IATPVPKTIVATGPTVPTNTVKEESVTETEAPKPVKSEEKVQSHGV D KANEVTKSDESSK
pFtf_Smut	---PSQKSINNTKETKKAHHVSTE-----KKQK-----K
pFtf_Bsub	-----
pFtf_Bamy	-----
pFtf_Bste	-----
pFtf_Eamy	-----
pFtf_Zmob	-----
pFtf_Adia	-----
pFtf_Ssal	GNNTKVA AKLATTPKTPSDSEGSNSNILSILATIFAAIASLALLGYGLVTGKIHL P KK
pFtf_Smut	GN-----SFFAALLALFSAFCV SIGFK-----
pFtf_Bsub	-----
pFtf_Bamy	-----
pFtf_Bste	-----
pFtf_Eamy	-----
pFtf_Zmob	-----
pFtf_Adia	-----

is highly conserved within the inulin producing Ftf of *S. mutans* and the levan-synthesising Ftf of *S. salivarius* (Fig 4.3). A highly conserved domain was also found to surround D³¹². This domain has been designated the 'sucrose box' (Sato and Kuramitsu, 1988; Fig 4.3) and is also found in enzymes which catalyse the transfer of fructose from sucrose such as the sucrose-6-phosphate hydrolases which otherwise show little homology to the Ftfs. It was originally thought that D³¹² might act as the nucleophile involved in the formation of a covalent fructosyl-enzyme intermediate. However, recent site-directed mutagenic studies by Song and Jacques (1999b) have shown that this region is in fact important in maintaining the structural integrity of the Ftf. Their analysis strongly indicates that a β -turn initiated at G³⁰⁵ is stabilised by the surrounding aspartic acid, serine, threonine and asparagine residues (Song and Jacques, 1999b; Mathews and van Holde, 1990). Interestingly, this amino acid does not appear to be conserved in *Z. mobilis* or in *A. diazotrophicus*, thus supporting this contention (Fig 4.3). In contrast, D³⁹⁷ is present in all Ftfs as part of an RDP motif (Fig 4.3). Site-directed mutagenesis suggests that this aspartic acid residue is the actual nucleophile responsible for catalysis (Song and Jacques, 1999b; Batista *et al.*, 1999).

4.3.3 The *S. salivarius* Ftf Contains a Novel Proline-rich C-terminal Region

A prime objective of this study was to define sequences necessary for attachment of the Ftf of *S. salivarius* to the cell surface. The Ftf sequence was therefore inspected for domains not present in the extracellular Ftf of *S. mutans* and the SacBs of the two *Bacillus* species. Multiple alignment of the deduced amino acid sequences (Fig 4.3) showed that the *S. salivarius* enzyme possessed a C-terminal domain which appeared to be unique, as homologous domains were not present in the other enzymes. This unique domain was proline-rich and terminated in a hydrophobic domain (Fig. 4.4). These features characteristic of plasma membrane anchoring and cell-wall-spanning regions of surface-bound polypeptides of Gram-positive origin (Fischetti *et al.*, 1991; Hansson *et al.*, 1992).

With a view to determining whether this domain was the product of evolutionary divergence or possibly the fusion of two sequences of disparate origin, the sequence of the C-terminal 240 amino-acids of the Ftf of *S. salivarius* was used to search protein sequence databases. This search failed to reveal any similarity with known Ftf or SacB sequences, but revealed homology to the product of the *prgC* gene from the plasmid, pCF10, of *Enterococcus faecalis* (Fig 4.5a). This gene codes for a pheromone-responsive protein of unknown function that is thought to be bound to the cell surface (Kao *et al.*, 1991). The Ftf of *S. salivarius* may therefore be the fusion of an enzymatic domain with a cell-wall anchoring domain. Interestingly, the proline-rich domain of the Ftf of *S. salivarius* also

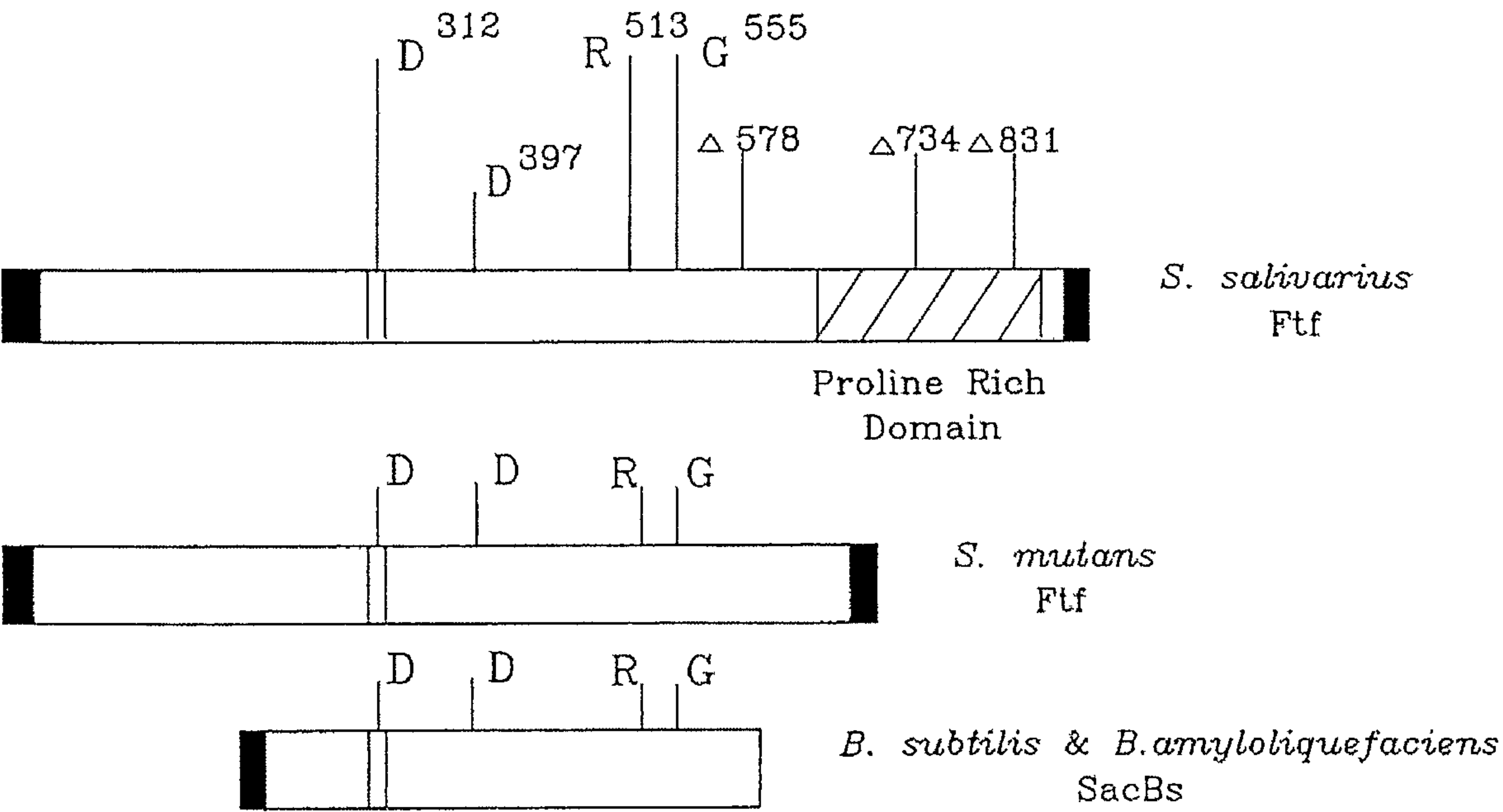


Figure 4.4 Domain structure of the Ftfs of *S. salivarius* and *S. mutans* compared with that of the SacBs of *B. subtilis* and *B. amyloliquefaciens*. Key amino acids involved in the structure and function of the SacB of *B. subtilis* are shown, as is the "sucrose box" surrounding D³¹² (Sato and Kuramitsu, 1988). The extent of the C-terminal deletion mutants in the Ftf of *S. salivarius* are also indicated (Δ⁸³¹, Δ⁷³⁴, Δ⁵⁷⁹). The shaded area represents the proline-rich domain and the solid blocks the N-terminal signal sequences and the C-terminal hydrophobic regions.

appeared to exhibit homology to the human proline-rich protein PRP-4 (Fig 4.5b) (Maeda *et al.*, 1985). However, as PRP-4 contains 82 proline residues out of a total of 248, one must view this alignment with some degree of scepticism. However, should the Ftf of *S. salivarius* represent a fusion protein derived from a eukaryotic source, this would be of particular evolutionary interest.

4.3.4 The C-terminus of the Ftf of *S. salivarius*

4.3.4.1 Cloning of Deletion Mutated *ftf* Genes into the Shuttle Vector pVA838

At the time of this study, no method for the transformation of *S. salivarius* ATCC 25975 was available. Consequently, the naturally transformable *S. gordonii* LGR2 was chosen as a model system for the expression of the *ftf* gene and various mutated forms..

Standard exonuclease III digestion (2.4.12) of the *ftf* gene was used to create a set of C-terminal deletion mutants of the *ftf* of *S. salivarius*. The phagemid pKRK104 (Fig 3.1; Table 2.1) containing the complete *ftf* gene was digested with *Xba*I and *Apa*I before treatment with exonuclease III. A set of three phagemids was constructed in this way. The first of these, pKRK1969, contained the intact *ftf* gene, but with 346 bp deleted from the downstream region. The second, pKRK1914, expressed a Ftf truncated at amino acid number 831 and the third, pKRK1816, expressed a Ftf truncated at amino acid number 734. A fourth phagemid, pKRK1661 expressing a Ftf deleted to amino acid number 578 was created by a simple *Hind*III digestion and re-ligation of pKRK104 (Table 2.1).

E. coli transformed with these phagemids expressed Ftf activities that hydrolysed sucrose and formed fructan except in the case of cells harbouring pKRK1661 where neither activity was observed. It is perhaps worth noting that if the equivalent deletion had been applied to the SacBs of the bacilli it would represent only a 23 amino acid deletion of the C-terminus, though in a region where a high degree of identity exists between all four enzymes. In order to study the effect of the C-terminal deletions on the localisation of the Ftf, the *ftf* gene and the various 3'OH deletion mutants were cloned into the *E. coli* - *Streptococcus* shuttle vector, pVA838. The phagemids, pKRK1969, pKRK1914 and pKRK1816 were each digested with *Eco*RI and *Sph*I to excise fragments containing the *ftf* gene or the truncated genes. These fragments were then cloned into pVA838 that had been digested with *Eco*RI and *Sph*I. This procedure gave rise to three new plasmids, pKRK2969, pKRK2914 and pKRK2816 (Table 2.1; Fig. 4.6).

4.3.4.2 Expression of *S. salivarius* Ftf Activity in *S. gordonii* LGR2

When plasmids pKRK2969, pKRK2914 and pKRK2816 were transformed into *S. gordonii* LGR2 significant recombination of the plasmids was observed giving rise to a high frequency of Ftf⁺, Em^r clones. Ftf⁺ transformants were therefore selected on BHI plates supplemented with 3% (w/v) raffinose and 40 µg erythromycin ml⁻¹. *S. gordonii* containing plasmids expressing Ftf activity produced colonies with large amounts of extracellular fructan and thus were easily distinguishable from transformants carrying re-arranged plasmids. Plasmids were extracted from fructan-producing transformants and characterised by restriction digestion in order to ensure that no rearrangements had taken place. Stable clones containing intact pVA838, pKRK2969, pKRK2914 and pKRK2816 were used in subsequent experiments.

4.3.4.3 Localisation of Mutated Ftf Activities in *S. gordonii*

S. salivarius Ftf activity was cell-bound only when the intact *ftf* gene was expressed in *S. gordonii*. *S. gordonii* carrying 3'OH truncated *ftf* genes on either pKRK2914 or pKRK2816 secreted 90-95% and 89-96% of their Ftf activity respectively (Table 4.2). Washed cells of *S. gordonii* carrying pKRK2969 released their cell-bound Ftf activity when incubated at 37°C for 20 min in the presence of 5 mM sucrose (Table 4.2). This was not the case in the absence of sucrose nor in the presence of its component sugars, glucose or fructose. This suggested that the mode of attachment of *S. salivarius* Ftf expressed in *S. gordonii* LGR2 may be similar to that in its native host, *S. salivarius* ATCC 25975 (Milward and Jacques, 1990).

4.4 DISCUSSION

The Ftf of *S. salivarius* possessed a unique domain not found in the Ftf of *S. mutans* or the SacBs of *B. subtilis* and *B. amyloliquefaciens*. This domain was at the C-terminus where a proline-rich region was located N-terminal to a hydrophobic domain. This juxtaposition of terminal domains is characteristic of a number of surface located proteins of Gram-positive cocci of which the M protein of *Streptococcus pyogenes* is a classic example (Fischetti *et al.*, 1991). Such surface proteins were originally considered to be anchored in the membrane by their hydrophobic C-termini with the region rich in proline and polar residues spanning the cell wall to give rise to a surface location of their N-terminal domains. However, these cell-bound proteins contain a pentapeptide with the consensus sequence, LPXTG, approximately nine amino acids N-terminal from the C-terminal hydrophobic domain (Fischetti *et al.*, 1990; Fischetti *et al.*, 1991). Evidence suggests that this pentapeptide is essential for the attachment of these proteins to the cell wall matrix following a C-terminal processing event rather than as

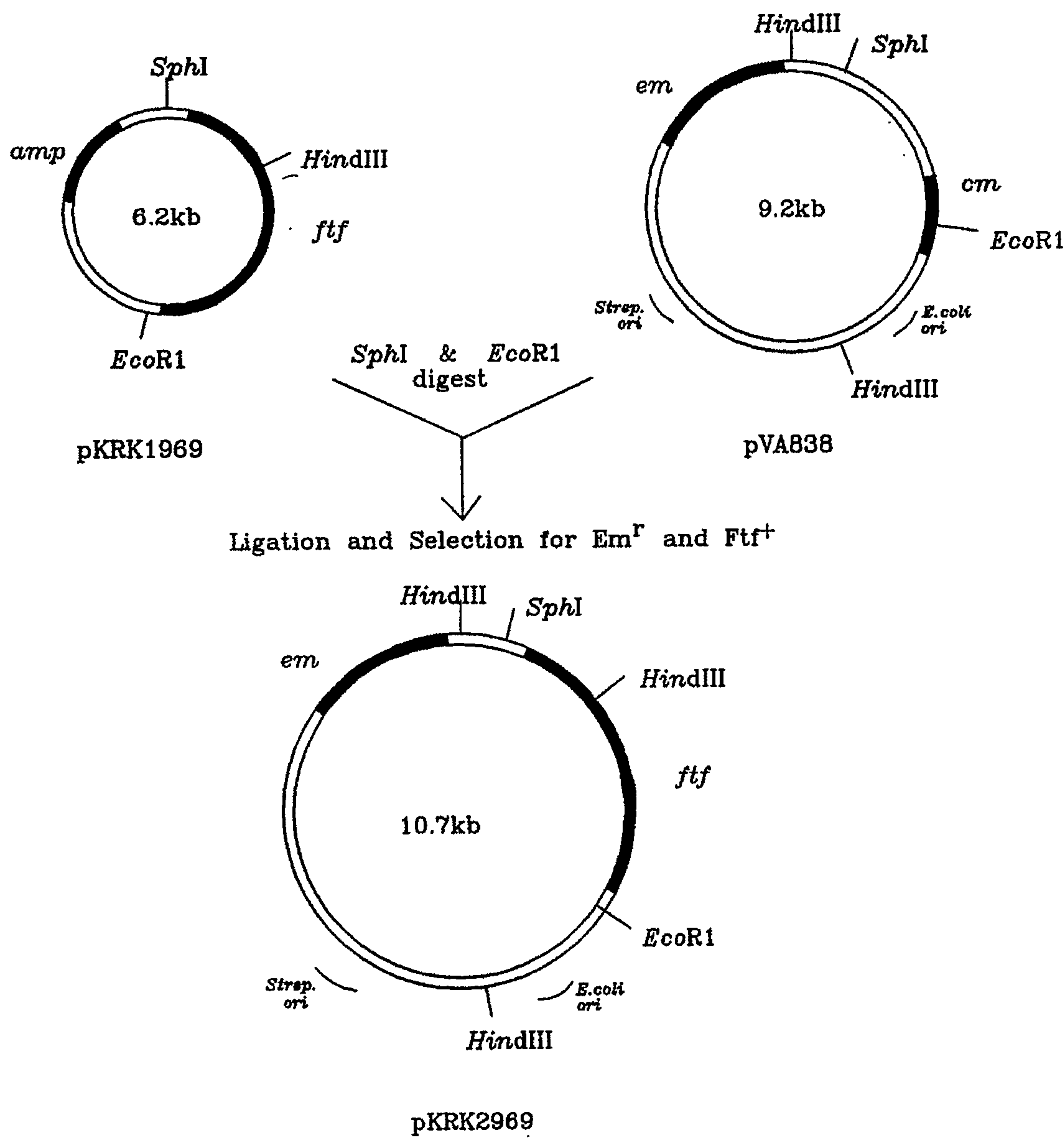


Figure 4.6 Construction of the *E. coli*/streptococcal shuttle vector pKRK2969 from pKRK1969 and pVA838. Similar strategies were used to construct pKRK2914 and pKRK2816 from pKRK1914 and pKRK1816 respectively which contained 3'OH deletions of the *ftf* gene.

Table 4.2. Cellular location of Ftf activity prior to and following incubation of washed cells with sucrose^a.

Bacterium	Plasmid	Extracellular Ftf Activity (%)	
		Cell Culture ^b	Sucrose Incubation ^c
<i>S. salivarius</i>	NONE	17 (14)	80 (80)
<i>S. gordonii</i>	NONE	0 (0)	- (-)
<i>S. gordonii</i>	pVA838	0 (0)	- (-)
<i>S. gordonii</i>	pKRK2969	48 (35)	71 (76)
<i>S. gordonii</i>	pKRK2914	90 (95)	- (-)
<i>S. gordonii</i>	pKRK2816	89 (96)	- (-)

^a Ftf activity for *S. salivarius* was within the range of 580-660 mU (mg dry wt)⁻¹ and for *S. gordonii*, transformed with the pKRK2969, pKRK2914 and pKRK2816, was 1220-1720 mU (mg dry wt)⁻¹, 700-1360 mU (mg dry wt)⁻¹ and 610-920 mU (mg dry wt)⁻¹, respectively.

^b Cells cultured in THBS for 16 h, were harvested, washed and resuspended as described in the text (2.3) and the amount of Ftf activity in the supernatant compared with that bound to the cell (2.5.2). The values in parentheses give the equivalent percentage values for cells grown in SDM (2.3).

^c Washed cells were incubated with 5 mM sucrose at 37⁰C for 20 min and the amount of Ftf activity released from the cell determined.

a motif for plasma membrane attachment (Schneewind *et al.*, 1992, 1993; Navarre and Schneewind, 1999). As this pentapeptide was not found in the Ftf of *S. salivarius* it was reasonable to assume that this enzyme was not attached to the peptidoglycan matrix of the cell wall, but was anchored to the plasma membrane by way of its C-terminal hydrophobic domain as envisaged for surface proteins of staphylococci that do not possess the LPXTG motif (Hansson *et al.*, 1992). Expression of native and C-terminally truncated Ftfs in *S. gordonii* were consistent with this view as only the native Ftf was initially bound to the cell; all C-terminal mutant enzymes being secreted into the culture fluid.

The only other known mechanism for binding cell surface proteins to Gram-positive bacteria is by way of a lipid anchor in the form of a *N*-acyl diglyceride modification of a cysteine residue C-terminal to the signal sequence that results in the formation of a lipoprotein (Pugsley and Schwartz, 1985; Wu and Hayashi, 1986; Sutcliffe *et al.*, 1993). However, no cysteine residue (which is a relatively rare amino acid in oral streptococcal proteins) is present in the Ftf of *S. salivarius*.

In conclusion, the results of this chapter not only strongly supported the idea that the C-terminal region of the Ftf of *S. salivarius* was involved in the surface location of the enzyme, but also showed that in *S. gordonii*, the processing event leading to its release and subsequent secretion from the cell required the enzyme's substrate, sucrose, as it did in its native host, *S. salivarius*. In light of this preliminary data using random exonuclease deletions, it was decided to determine more rigorously the role the various domains of the C-terminus of the Ftf of *S. salivarius* played in the anchoring the enzyme to the cell surface. The approach taken was to use site-directed mutagenesis to construct specifically defined mutated *ftf* genes. In order to achieve this, however, it soon became apparent that it would be necessary to develop a novel method for site-directed mutagenesis. The development and application of this method is discussed in the next chapter (Chapter 5).

CHAPTER 5

DEVELOPMENT OF A TECHNIQUE FOR MULTIPLE SITE-DIRECTED MUTAGENESIS OF THE *ftf* GENE CONTAINING PALINDROMIC SEQUENCES¹

5.1 INTRODUCTION

The previous exonuclease deletion studies of the *ftf* gene indicated that binding of Ftf to the cell surface depended upon the presence of both a proline-rich extension and a hydrophobic domain at the C-terminus (Chapter 4). In order to distinguish the role of these two domains in surface attachment, attempts were made to mutate the *ftf* gene by *in vitro* site-directed mutagenesis. Mutagenic oligonucleotides were designed to introduce a series of restriction sites within the proline-rich wall-associated domain of the *ftf* gene, allowing serial deletion of sections in this region while maintaining the hydrophobic tail. Another oligonucleotide allowed the introduction of a stop codon prior to the hydrophobic domain resulting in a truncated form of the protein. However, numerous early attempts to mutagenize the *ftf* gene using standard protocols were unsuccessful as they resulted in a series of deletions of the gene. These deletions appeared to commence at points within the *ftf* gene where there were palindromic sequences that were capable of forming closed loop structures which acted as terminators under conditions of mutagenesis. To overcome this problem, two modified mutagenic techniques were developed.

5.2 METHODS

5.2.1 Chemicals and Enzymes

The TransformerTM Site-Directed Mutagenesis Kit from Clontech was used as the basis for site-directed mutagenesis (2.1). Other chemicals and enzymes were obtained from the sources also listed in Section 2.1.

5.2.2 Bacterial Strains, Phagemids, Bacteriophage and Growth Conditions

Escherichia coli strains BMH 71-18 *mut* S, CJ236 and NM522 (Table 2.1) grown at 37 °C in LB medium supplemented with ampicillin (2.3) were used as hosts for the various

¹ Some of the results of this chapter have been published – Rathsam and Jacques, *FEMS Microbiol. Letts.* 153: 447-453., 1997.

site-directed mutations of phagemid pKRK1969 (4.3.4.1; Table 2.1). Single stranded DNA was produced from pKRK1969 by superinfection with M13K07 (2.4.5; Table 2.1).

5.2.3 DNA Manipulations and Mutagenesis

The amplification and CsCl purification of phagemid DNA were as previously described (2.4.2; 2.4.4). Initial failed attempts at single-stranded and double-stranded site-directed mutagenesis with either a uricilated or a non-uricilated template of pKRK1969 were based on the three procedures described by Sambrook *et al.* (1989), Slilaty *et al.* (1990), and the instructions supplied with the Transformer™ Site-Directed Mutagenesis Kit. In all cases the ability of all oligonucleotides to anneal to and act as mutagenic primers for the pKRK1969 template was investigated by determining whether they could act as primers for sequencing (2.4.16).

Table 5.1 Mutagenic oligonucleotides

Primer	New Restriction Site	Oligonucleotide*
PrS	<i>Mlu</i> I → <i>Bsr</i> I	TCCGTCGTGAGGAG <u>ACTGGT</u> GGCATGCAAGCTT
Pr1	Stop Codon, <i>Dra</i> I	CAAGTGATTCAGAAG <u>TTTAAA</u> ACAGCAAC
Pr2	<i>Bam</i> HI	GACTGGATTTGGG <u>ATCC</u> ATCTAGTGAAAGCTTG
Pr3	<i>Bam</i> HI	TCCTAAGACACCAAG <u>GGATCC</u> AGAAGGTTCAAACA
Pr4	<i>Bam</i> HI	GACTGTAAGTGT <u>TAAGGATC</u> CTAAAGACAAGAAGGC
Pr5	<i>Bam</i> HI	ATCAACCATCAGC <u>GGATCC</u> AGAAGATAAACCAAC
T3	---	TTAACCCTCACTAA

* The changed nucleotides are shown in bold and the introduced restriction site is underlined.

The successfully developed methods of site-directed mutagenesis used the solutions and enzymes provided in the Transformer™ Site-Directed Mutagenesis Kit with a number of modifications to the manufacturer's instructions. Mutagenic oligonucleotides of 29-36 bp incorporating two or three base changes (Table 5.1) were phosphorylated before annealing.

Phosphorylation was accomplished by incubating 2 µg of primer with 3 µl of 10x kinase buffer, 1 µl of T4 polynucleotide kinase (10 units) and 1 µl of 30 mM ATP in a total volume of 30 µl for 60 min at 37°C. After 60 min, the T4 polynucleotide kinase was inactivated by heating at 68°C. The alkaline method of denaturation using 2 µg of double stranded DNA in a final volume of 30 µl of H₂O followed by the annealing reaction described in the manual for the Transformer™ Site-Directed Mutagenesis Kit was found to be the most efficient. For the annealing reaction a fifth of the denatured template (6 µl), 1 µl of the phosphorylated selection primer(s), 2 µl of the appropriate mutagenic primer and 2 µl of 10x annealing buffer was incubated at 37°C for 20 min followed by 10 min at room temperature. Synthesis of the mutated DNA strand was subsequently carried out in a final volume of 30 µl. As the *fitf* gene is AT rich, an additional 1 µl of a 10 mM solution of both dATP and dTTP was added to the 3 µl of synthesis buffer containing 1 µl of T4 DNA polymerase and 1 µl of T4 DNA ligase as supplied in the Transformer™ kit. This reaction mixture was incubated at 4°C for 10 min, a further 10 min at room temperature (18-22°C) and then at 37°C for 2 h, after which time one of the following two alternative procedures were followed.

5.2.3.1 The T7 Polymerase Method

T7 DNA polymerase (2 units, 1µl) was added and the reaction mixture incubated at 37°C for 10-15 min. The reaction was stopped by heating at 68°C for 10 min. T4 DNA ligase (2U; 1µl) and 1 mM ATP were added and the reaction mixture was incubated at 16°C for 2 h. This ligation mixture was heat inactivated at 68°C for 10 min before precipitating the DNA with 75% (v/v) ethanol and 10% (v/v) 3 M sodium acetate. The DNA was reconstituted to a final volume of 30 µl in an appropriate buffer for digestion with the unique selection restriction endonuclease, *Mlu*I.

5.2.3.2 The Vent Polymerase Method

The DNA was first precipitated and dried as in Section 5.2.3.1 above prior to being re-dissolved in 85 µl of H₂O, containing 10 µl of 10x Vent polymerization buffer (supplied by the manufacturer), 5 µL of 100 mM MgSO₄ and 1 µL of 20 mM dNTP solution. One unit of Vent DNA polymerase was added and the reaction mixture was incubated at 74°C for at least 3 min (the incubation time was determined by the average size of the deletions observed [see below] and the fact that Vent DNA polymerase polymerizes at a rate of about 1 kbp min⁻¹). The DNA was precipitated as above, dried and re-dissolved for ligation in a final

volume of 30 μ l using the appropriate buffer for the selection restriction endonuclease, *Mlu*I. Prior to restriction, the ligation mixture was heat inactivated at 68⁰C for 10 min.

In all cases the putative formation of the required site-directed mutated *ftf* gene was confirmed by sequencing (2.4.16) through the mutated restriction site(s) using appropriate oligonucleotide primers (Table 4.1).

5.2.4 Transformation and Selection

Competent *E. coli* BMH 71-18 *mutS* cells defective in mismatch repair were electroporated (Bio-Rad Gene Pulser; 1.6kV, 200W, 25mA) using 3 μ l of heat-inactivated DNA solution that had been restricted with *Mlu*I. The recovery and secondary selection steps were carried out according to the protocol described in the Clontech TransformerTM Site-Directed Mutagenesis manual. In essence, plasmid DNA was prepared from the mixed bacterial population, re-digested with *Mlu*I and transformed yet again into competent *E. coli* BMH 71-18 *mutS*. The final transformation using the DNA that had been thoroughly digested with *Mlu*I resulted in a high recovery of the desired mutated plasmid which was selected according to its unique *Dra*I or *Bam*H1 digestion pattern.

5.3 RESULTS

The aim of the site-directed mutagenic experiments was to introduce unique restriction sites within the 3'-OH region of the *ftf* gene in order to construct unique deletions within the gene such that Ftf proteins with defined truncated C-termini could be expressed. The position of these primers relative to the *ftf* gene and the domains coded by the gene are summarized in Figure 5.1.

5.3.1 Formation of Deletions Using Standard Site-Directed Mutagenesis Protocols

Prior to any attempt at mutagenesis of PKRK1696, the ability of the site-directed mutagenic primers to individually anneal to and act as primers for sequencing reactions was checked (2.4.16). In all cases the designed primers specifically bound to only one site on pKRK1969 and the primers and template were found to be of sufficient purity to provide clear sequencing reactions. First attempts at site-directed mutagenesis concentrated on using the selective primer, PrS, which is used in the TransformerTM site-directed mutagenesis protocol to introduce a unique restriction site in order to reduce the frequency of wide-type clones (see Fig 5.8 below) together with Pr1 which introduces a stop codon and a *Dra*I site as well as the selective primer, PrS, in conjunction with Pr2 and Pr3 to introduce two unique *Bam*HI sites (Table 5.1). Various published methods of mutagenesis were attempted as described below

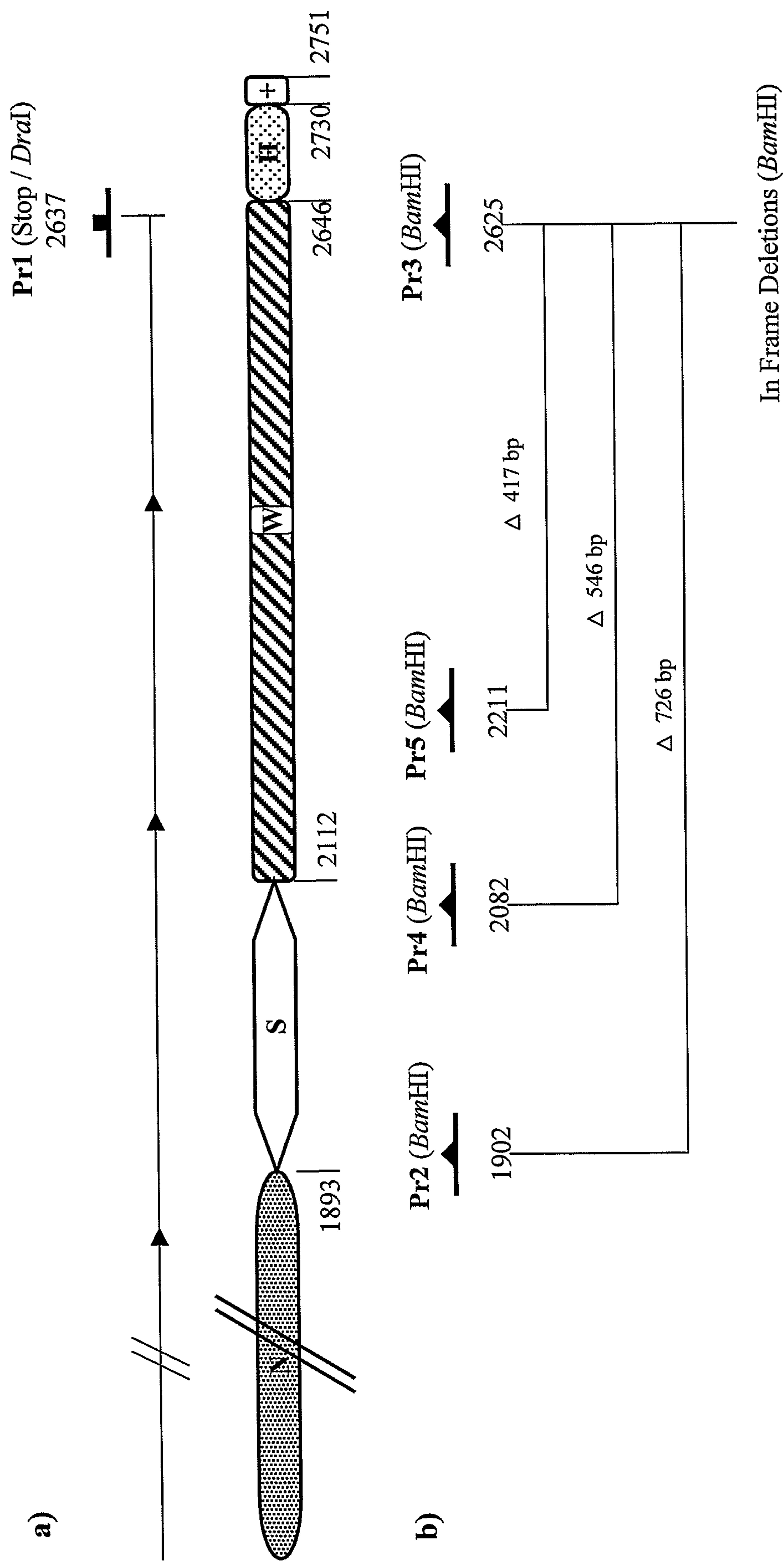


Figure 5.1 Position of mutagenic primers relative to the deduced amino acid sequence of the *ftf* gene. (a) Pr1 introduces a stop codon immediately prior to the hydrophobic tail. (b) Pr3 + Pr2, Pr3 + Pr4, Pr3 + Pr5 introduce *Bam*HI sites allowing for in-frame deletions. N, amino-terminus containing the catalytic domain; S, spacer region; W, wall-associated domain; H, hydrophobic domain; +, positively charged C-terminus. Numbers refer to base pairs within the *ftf* gene

(Sections 5.3.1.1; 5.3.1.2).

5.3.1.1 Single-Stranded Mutagenesis

Single-stranded and uracil-containing single-stranded DNA of phagemid pKRK1969 was produced by superinfection of *E. coli* NM522 hosting pKRK1969 with bacteriophage M13K07 (Table 2.1). Although the yield of single-stranded DNA varied (possibly due to the large size of the phagemid) several batches of single-stranded DNA from pKRK1969 were eventually obtained. Only batches of single stranded DNA that gave a clear sequencing ladder were used as templates for site-directed mutagenesis. Initially no transformants were obtained with the primers PrS and Pr1 irrespective of whether T4 DNA polymerase or Klenow were used in the extension reactions for site-directed mutagenesis. However, later attempts yielded transformants which on analysis contained various deletions (Fig 5.2). As all the mutagenic primers were clustered within an approximate 1 kbp region of a 6.2 kbp phagemid, an additional commercial primer, primer T3, that was homologous to the T3 promoter region of the pIBI31 vector and which lay approximately 2.3 kbp upstream of the other primers was also included (see Fig 4.1). While the inclusion of this primer in the reaction mixture altered the profile of the clones, slightly increasing the number of large oversized phagemids, a predominance of deleted mutants still existed (Fig 5.3).

5.3.1.2 Double-Stranded Mutagenesis

Double-stranded site-directed mutagenesis was also attempted on closed circular as well as linearised phagemid DNA according to the method described by Slilaty *et al.* (1990) as well as the protocol supplied with the Clontech TransformerTM site-directed mutagenesis kit. Both non-uracilated and uracilated DNA generated in *E. coli* strains NM522 and CJ236 were used (Table 2.1). Transformation and amplification in *E. coli* NM522 following attempted mutagenesis with primers PrS and Pr1 resulted in the recovery of either the native 6.2 kbp phagemid or a series of ampicillin resistant phagemids of 2.2 kbp, 2.3 kbp, 2.4 kbp or 3.4 kbp (Fig 5.4). Larger phagemids of the order of 8-10 kbp or larger were sometimes isolated. Mutagenesis was also attempted with the primers, PrS, Pr2 and Pr3, which if successful should have introduced two *Bam*HI sites into pKRK1696. As pKRK1696 has no *Bam*HI site and as many of the screened clones were undigested by *Bam*HI, this indicated that only wild-type or deletion mutations had been recovered (Fig 5.5).

5.3.2 Production and Mutagenesis of a 5.4 kbp Phagemid containing a Partial *ftf* Gene

As numerous attempts at mutagenesis of pKRK1696 resulted in deletions, a strategy

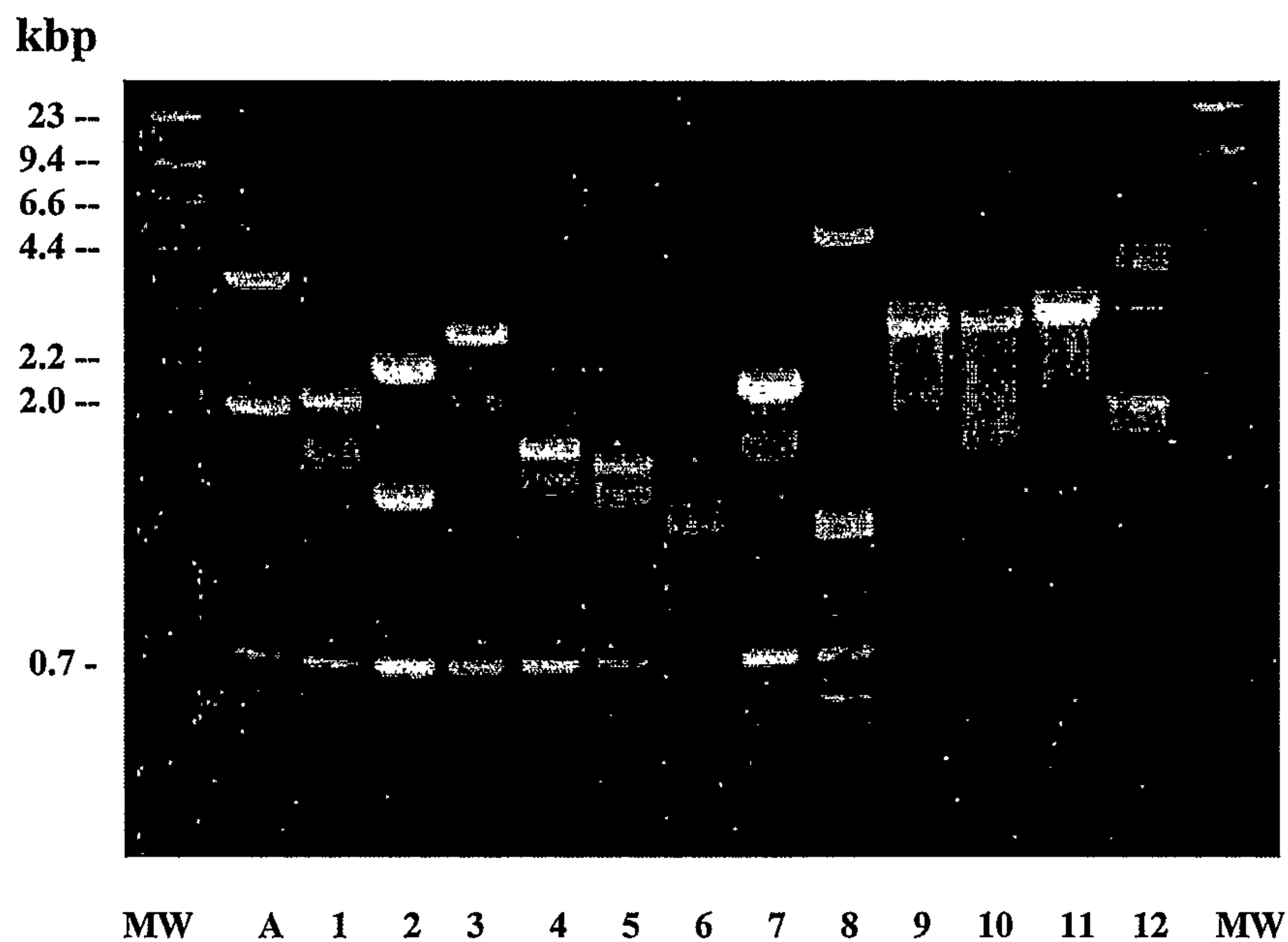


Figure 5.2 Gel electrophoresis showing *Dra*I digested phagemids produced from the mutagenesis of uricilated single-stranded pKRK1696 using primers PrS and Pr1. Lane A shows a standard pKRK1696 *Dra*I digest (3.5, 2.0, 0.7 kbp); Lanes 1 & 7 (2.0, 0.7 kbp); Lane 2 (2.2, 1.2, 0.7 kbp); Lane 3 (2.6, 0.7 kbp); Lane 4 (1.6, 0.7 kbp); Lane 5 (1.5, 0.7 kbp); Lanes 6 & 12 (uncut); Lane 8 (>6.2 kbp); Lane 9, 10 & 12 (2.6 kbp). MW, molecular weight markers.

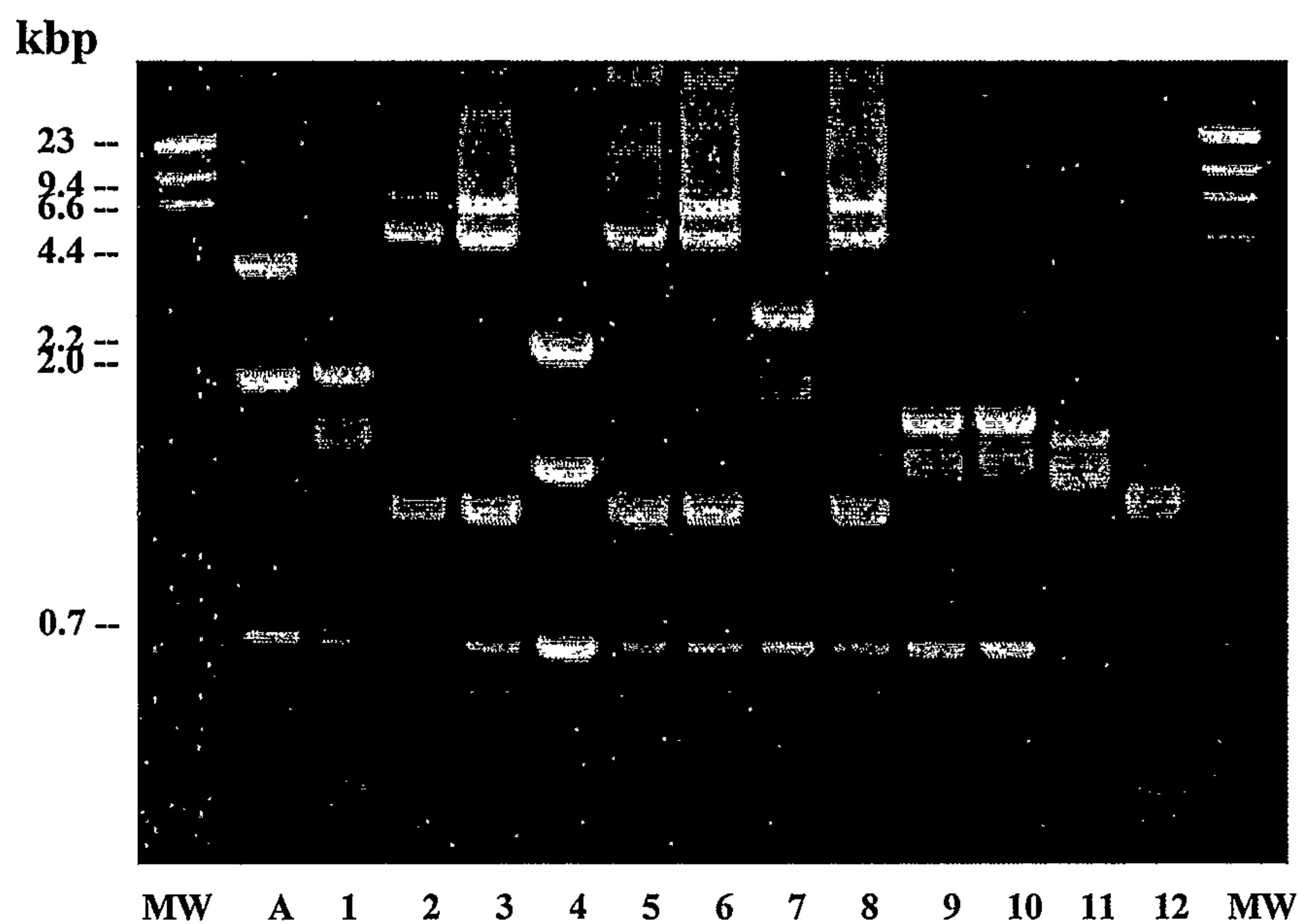


Figure 5.3 Gel electrophoresis showing *Dra*I digested phagemids produced from the mutagenesis of uricilated single-stranded pKRK1696 using primers PrS, Pr1 and T3. Lane A shows a standard pKRK1696 *Dra*I digest; Lane 1 (2.0, 0.7 kbp); Lanes 2, 3, 5, 6 & 8 (>6.2 kbp); Lane 4 (2.2, 1.2, 0.7 kbp); Lane 7 (2.6, 0.7 kbp); Lanes 9 & 10 (1.6, 0.7 kbp); Lane 11 (1.5, 0.7 kbp); Lane 12 (uncut). MW, molecular weight markers.

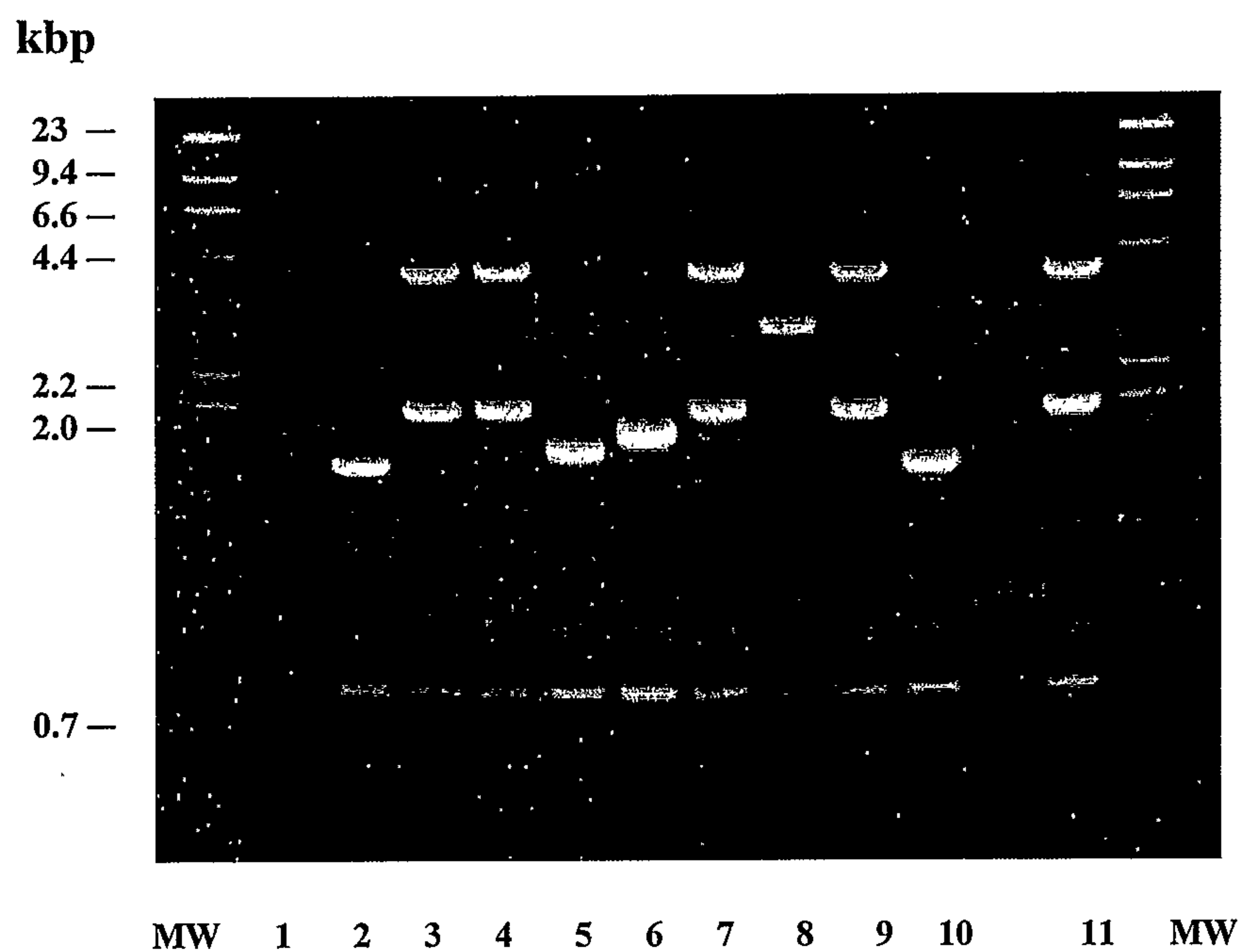


Figure 5.4 Gel electrophoresis showing *DraI* digested phagemids produced from the mutagenesis of uricilated double-stranded pKRK1696 using primers PrS and Pr1. Lanes 3, 4, 7, 9 & 11 are all wild-type pKRK1696 (3.5, 2.0, 0.7 kbp); Lane 1 (2.0, 0.7 kbp); Lanes 2 & 10 (1.5, 0.7 kbp); Lane 5 (1.6, 0.7 kbp);Lane 6 (1.7, 0.7 kbp); Lane 8 (2.6, 0.7 kbp). MW, molecular weight markers.

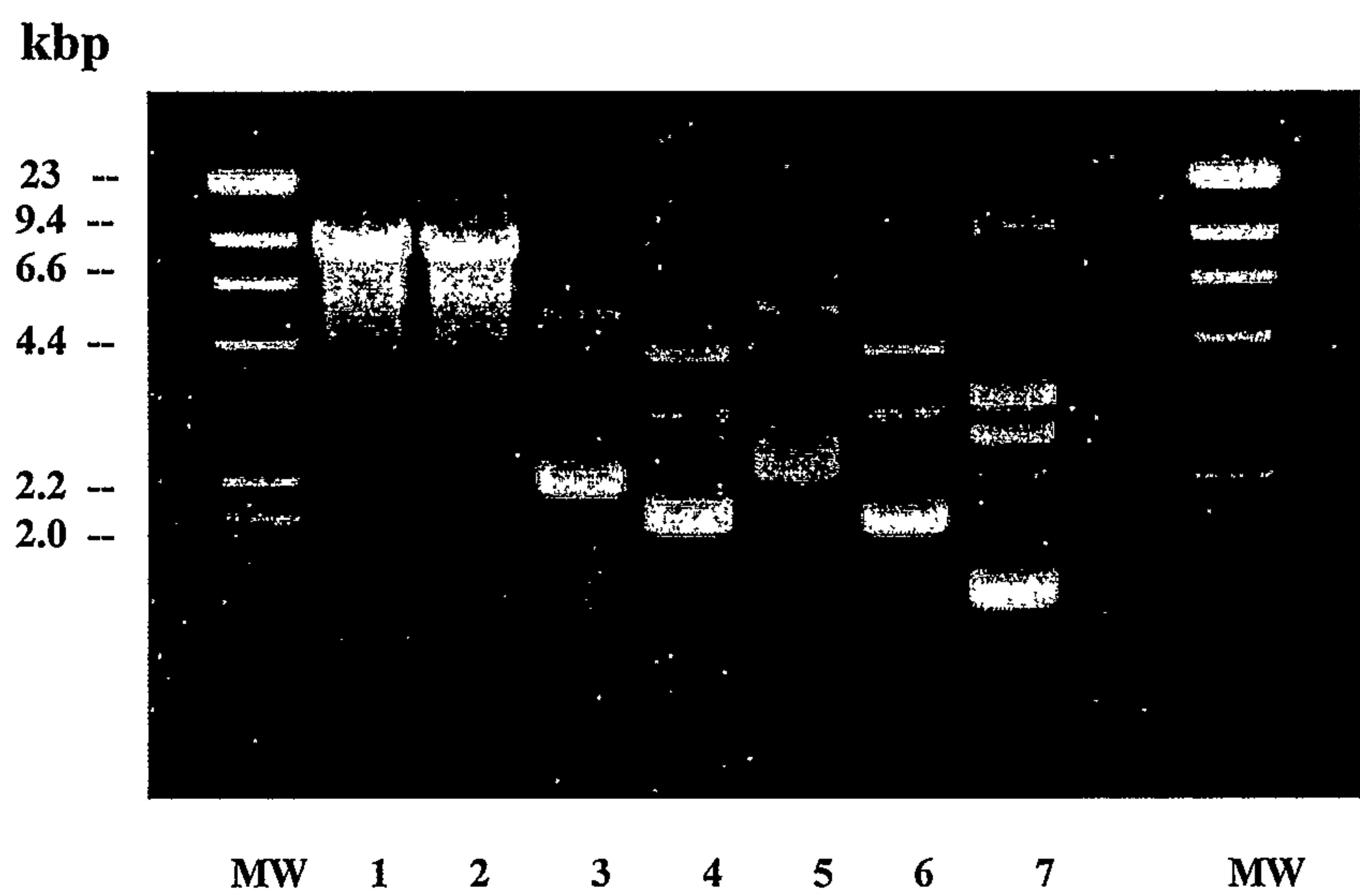


Figure 5.5 Gel electrophoresis of *Bam*HI digested phagemids produced from the mutagenesis of uricilated double-stranded pKRR1696 using primers, PrS, Pr2 and Pr3. Lanes 1 & 2 show large clones >8 kbp; Lanes 3-7 show undigested clones.

was devised to reduce the size of the phagemid on the assumption that its size might be responsible for the failure to obtain the required site-directed mutants. A 2.5 kbp *EcoRV* fragment of the 3.3 kbp *ftf* insert was re-cloned into the vector pIBI31, resulting in a deletion of 0.8 kbp from the 5'-OH-end of the *ftf* gene and the formation of pKRK3969 (Fig 5.6; Table 2.1). However, all attempts at both single and double-stranded mutagenesis of pKRK3969 resulted in deletions.

5.3.3 Examination of the Stem-Loop Termination Points and Mutagenesis of the *ftf* Gene by Modified Techniques

The consistent formation of a particular series of deleted plasmids seemed to indicate that the DNA polymerases were being interrupted during the extension reactions at specific sites within the phagemid and that recombination events were occurring to produce a series of smaller phagemids containing the *amp* gene. The presence of the 0.7 kbp fragment (which is derived from the pIBI31 vector during *DraI* digestion) in nearly all the clones screened, suggested that the deletions occurred at some point in the 3'-OH-end of the coding region within the *ftf* gene. Analysis of the *ftf* gene revealed the existence of numerous palindromic sequences which could theoretically form stem-loop structures after denaturation of the double-stranded phagemid, pKRK1969 (Fig 5.7). Termination of DNA polymerization by these structures when either T4 DNA polymerase or Klenow was used to extend the annealed mutagenic primers was considered a possibility. In fact, if extension was to proceed from the primers and through the 2.9 kbp vector unto the stem-loops, the deleted clones should be > 2.5 kbp, 4.4 kbp and 4.5 kbp. On analysis, the clones screened fall into two categories; the majority below 2.5 kbp (2.4, 2.3, 2.2 kbp) and below 4.4 kbp (4.1, 3.4, 3.1, 2.7 kbp).

Since sequencing reactions using T7 polymerase appeared to proceed unhindered and clearly polymerization at elevated temperatures should preclude formation of these secondary structures, T7 and Vent DNA polymerase were chosen to replace the T4 DNA polymerase. In these instances, the protocol and all other reagents that were supplied in the Clontech TransformerTM Site-Directed Mutagenesis kit were used. However, exchange of T4 DNA polymerase with either T7 DNA polymerase at 37⁰C or Vent DNA polymerase at 74⁰C failed to produce any site-directed mutants. Native pKRK1969 was generally recovered in low yield following transformation. A combination of either T4 and T7 DNA polymerase or T4 and Vent DNA polymerase was then tested. In both cases successful site-directed mutations within the *ftf* gene were obtained (Fig 5.8). Whilst the method employing the use of the Vent DNA polymerase produced proportionally more mutants, the T7 method allowed the direct formation of double site-directed mutations (Table 5.2).

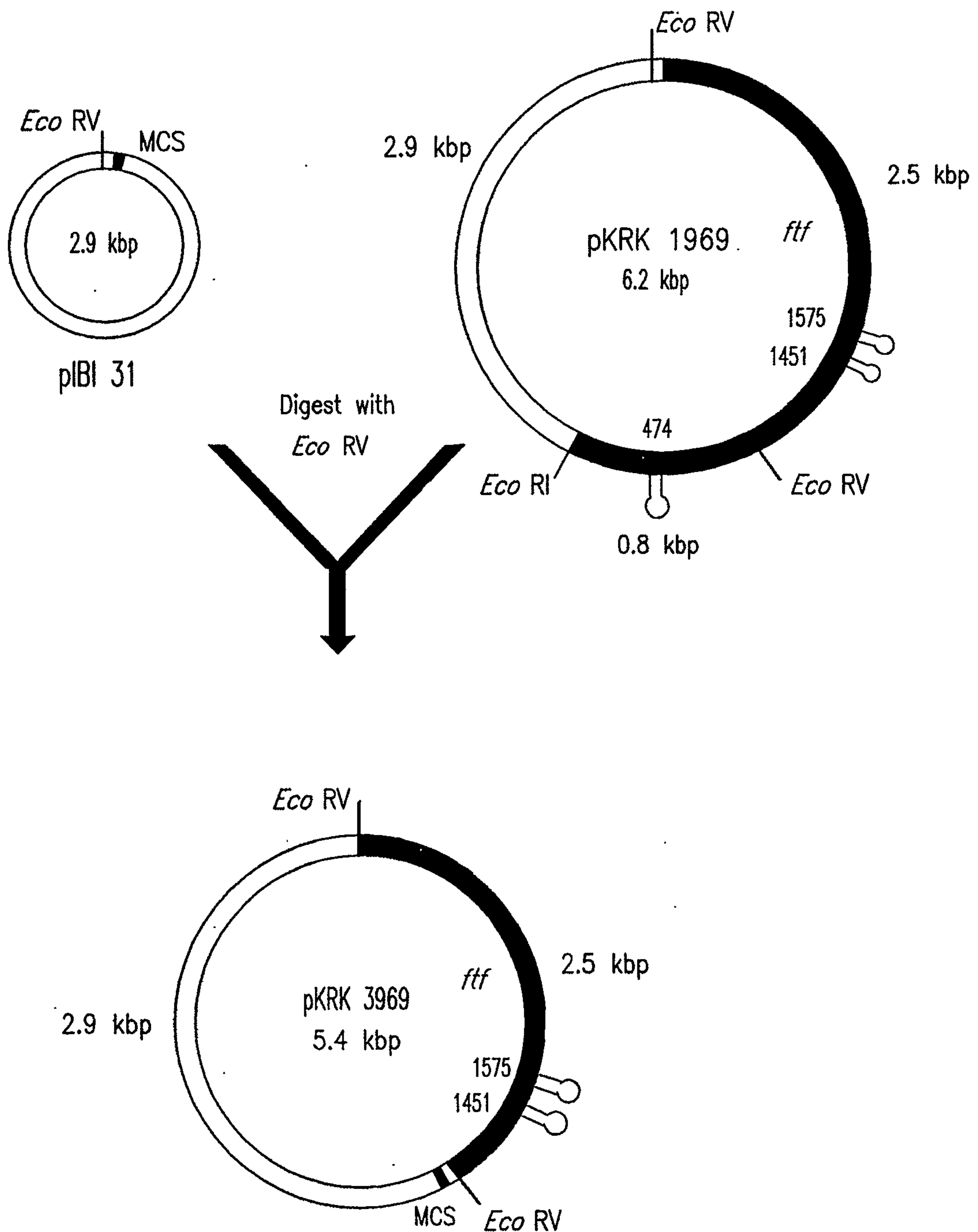


Figure 5.6 Construction of pKRK3969 from pKRK1969 and pIBI31. The shaded area shows the *S. salivarius* chromosomal insert containing the *ftf* gene which is transcribed in an anti-clockwise direction. The numbers next to the stem loop structures indicate the base pair number within the chromosomal insert where the structures commence. MCS, multiple cloning site.

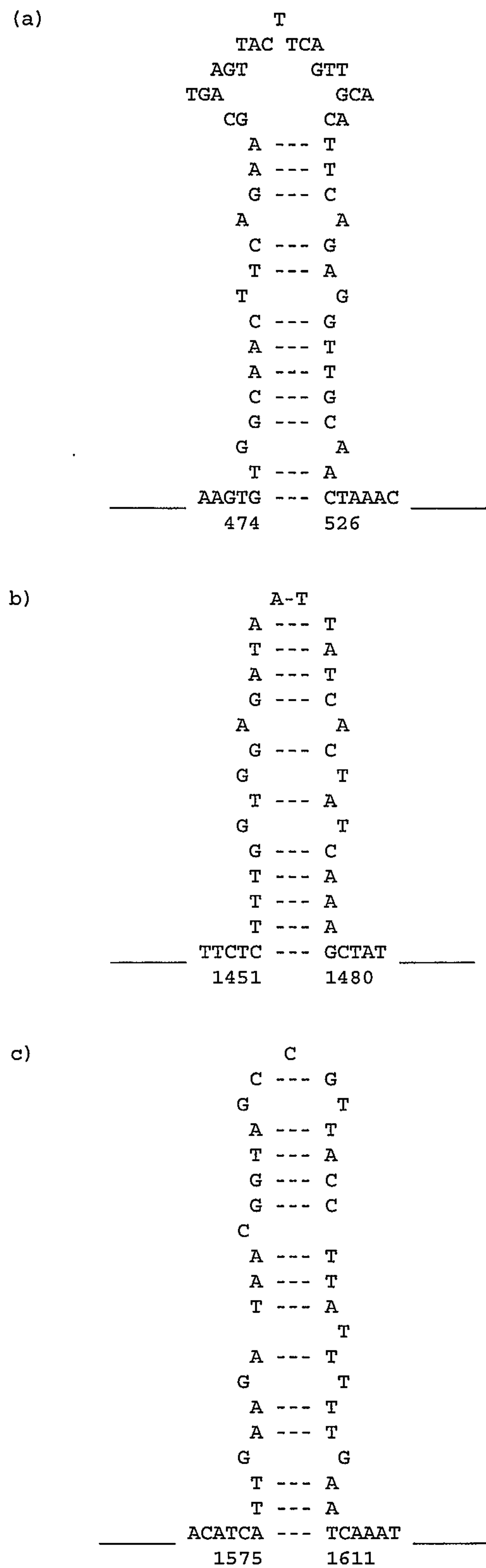


Figure 5.7 Diagrammatic representation of three major putative secondary structures found amongst many others within the *S. salivarius* chromosomal insert in pKRK1969 starting at, (a), 474 bases into the insert [80% homology]; (b), 1451 bases into the insert [80% homology] and (c), 1575 bases into the insert [82% homology].

Table 5.2 Frequency and number of colonies screened for successful site directed mutations.

Mutagenesis Protocol*	Number of Mutants/Colonies Screened					
	Pr1**	Pr2	Pr3	Pr2+Pr3	Pr2+Pr4	Pr2+Pr5
T7 Method	1/12	1/6	1/6	1/6	1/12	1/2
Vent Method	5/12	4/12	3/12	0/12	ND	ND

* For further details, see Sections 5.2.3.1 & 5.2.3.2..

** Mutagenic Primer(s) (see Table 5.1)

ND Method not attempted.

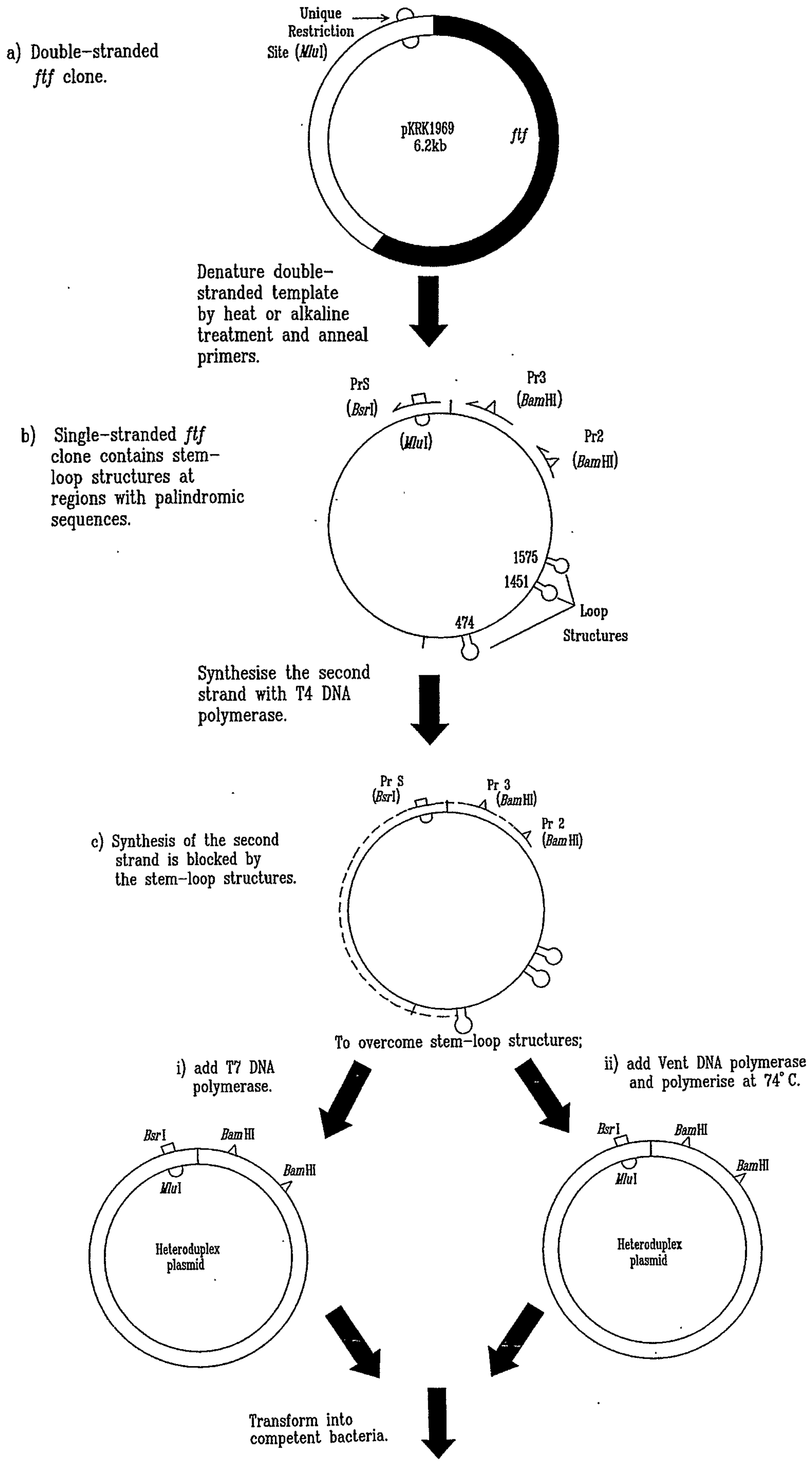
5.3.4 Sequencing of Mutated *ftf* Genes

All mutated *ftf* genes were screened by restriction digestion before confirmation by sequencing. The site-directed mutagenic phagemids created for subsequent engineering and expression of mutated Ftf proteins are as described in Fig. 5.1.

5.4 DISCUSSION

A primary requirement for site-directed mutagenesis is the provision of a suitable DNA template for annealing mutagenic primers. Standard protocols suggest subcloning the targeted segment of the gene into bacteriophage M13 or its derivatives in order to produce single-stranded DNA. If no suitable restriction sites are found, as was the case with the *ftf* gene, then it is recommended that restriction sites be introduced by site-directed mutagenesis. Clearly, this may prove problematic if the clone is not amenable to mutagenesis in the first place. Another disadvantage of this method is that multiple subcloning steps are required before a suitable clone is obtained. Large inserts cloned into M13 vectors may also render the construct unstable and prone to deletions. Careful handling may minimize this problem but will not completely eliminate it (Sambrook *et al.*, 1989). Phagemid vectors overcome both these problems. However, they are likely to produce low and inconsistent yields of single-stranded DNA (Zagursky and Berman, 1984).

This investigation demonstrated that the difficulties encountered with site-directed mutagenesis of the *ftf* gene were most likely due to the presence of palindromic sequences



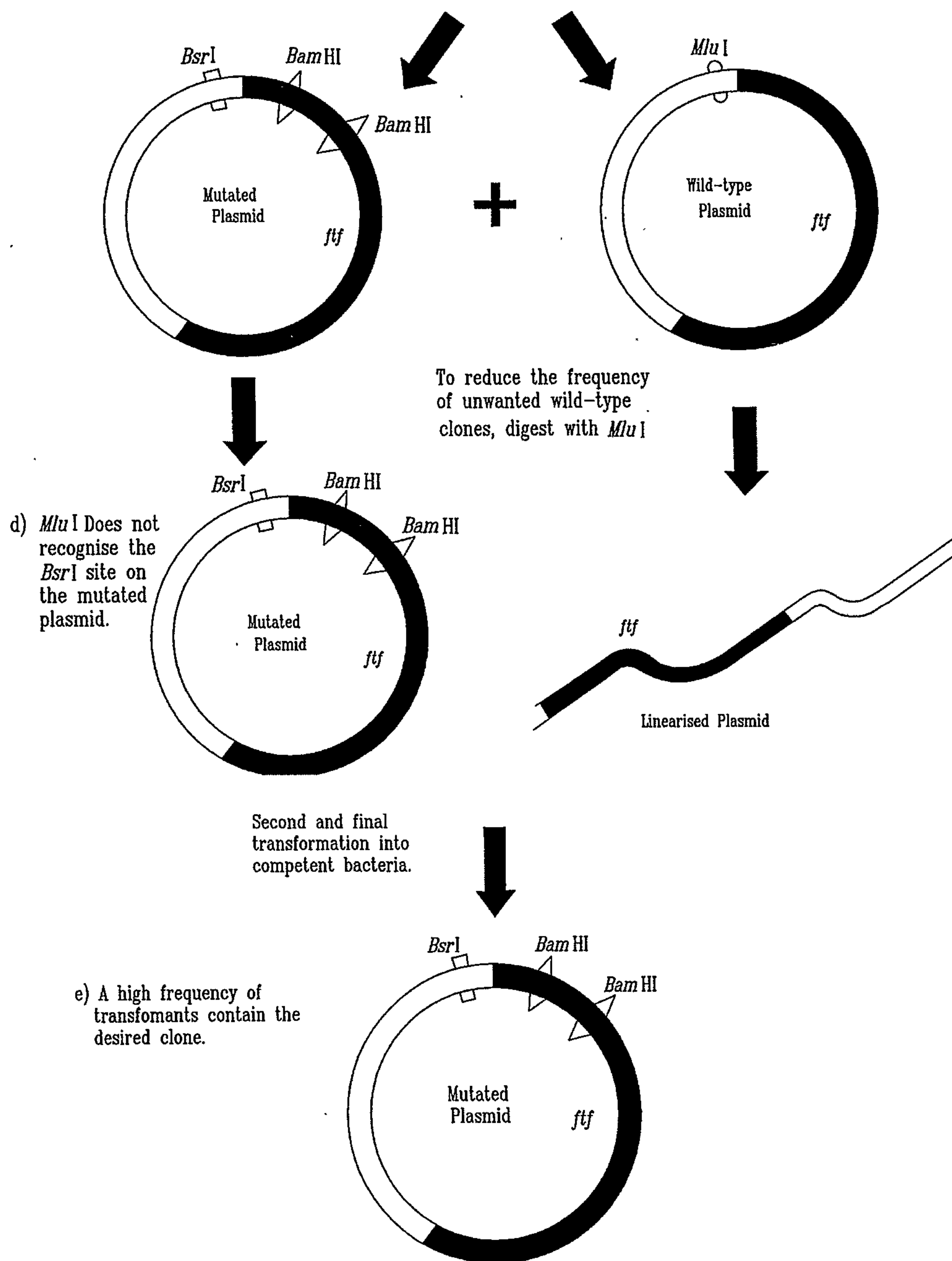


Figure 5.8 Strategy for *in-vitro* oligonucleotide mutagenesis of the *ftf* gene; (a) the double stranded template, pKRK1969 was denatured by alkaline treatment and the mutagenic primers annealed (b). The presence of putative stem-loop structures and their location within the *S. salivarius* chromosomal insert are shown. Synthesis of the second strand with T4 DNA polymerase was considered to proceed until blocked by the stem-loop structures within the *ftf* gene (c). To overcome the stem-loop structures either T7 DNA polymerase at 37°C or Vent DNA polymerase at 74°C was added then digested with *MluI* to reduce the frequency of unwanted native pKRK1969. The resultant digest was transformed into competent *E. coli*. Replication of the heteroduplex DNA (d) resulted in the formation of both mutated and native *ftf* alleles. These phagemids were digested once more with *MluI* and re-transformed into *E. coli* to obtain a high frequency of the desired mutated allele (e).

within the gene that formed stable secondary structures that disrupted the synthesis of DNA by T4 DNA polymerase or Klenow. Whilst formation of stem-loop structures has been previously reported to interfere with mutagenesis, these structures have only been associated with the target sequence, effectively reducing the efficiency of annealing of the mutagenic primers (Sambrook *et al.*, 1989). Why T7 DNA polymerase or Vent DNA polymerase alone was unable to completely polymerize the circular template without the prior use of T4 DNA polymerase remains unclear as both of these enzymes have been successfully used to produce site-directed mutations and to overcome stable secondary structures (Byrappa *et al.*, 1995; Cariello *et al.*, 1991). However in the case of modified T7 DNA polymerase, only a short fragment of dsDNA (389 bp) containing one potential stable secondary structure was amplified without any observed deletion (Cariello *et al.*, 1991). One possible explanation for our finding is that after extension by T4 DNA polymerase, the addition of T7 DNA polymerase allows the formation of a more stable complex that unfolds the secondary structures. The Vent DNA polymerase method that was developed relied on the elevated temperature of polymerization to reduce formation of hairpin loops in the template. At these temperatures the annealed mutagenic primers may also become unstable. The T4-extended oligonucleotides, however, would be expected to provide a much larger and far more stable primer from which the Vent DNA polymerase could extend.

Recently many polymerase chain reaction (PCR)-based site-directed mutagenesis methods have been devised. Two main advantages of these methods are that virtually any double-stranded vector may be used and, secondly, that there is a significant reduction in the number of multiple subcloning steps. Three current PCR-based methods will allow the introduction of multiple site-directed mutations (Shayiq and Black, 1994; Dwivedi *et al.*, 1994; Rouwendal *et al.*, 1993). The first involves the use of an asymmetric polymerase chain reaction (Shayiq and Black, 1994) with the entire process requiring three sequential PCR steps, two gel purifications and a cloning step. The second method, referred to as ‘multiple mutagenesis with a single selection primer’, requires two separate PCR steps as well as two digestion and four transformation steps for the introduction of two mutations (Dwivedi *et al.*, 1994). The third requires the ordered coupling of PCR generated fragments catalyzed by a thermostable DNA ligase. It is a multi-step process that is dependent upon the production of a ssDNA template (Rouwendal *et al.*, 1993). The multiple *in vitro* site-directed mutagenesis method described here can be performed on a single dsDNA template without any subcloning or PCR steps, and with only two transformations. This is not only time-saving but also reduces the chance of errors being introduced by the PCR reaction. The method also allows

multiple site-directed mutagenesis of large genes such as the *ftf* gene of *S. salivarius* containing palindromic sequences that may form DNA-polymerase-inhibiting secondary structures to be rapidly and effectively accomplished. As a result, all of the site-directed mutations were successfully engineered with the desired modifications for use in producing the required deletions within the *ftf* gene and subsequent expression of Ftf proteins possessing various C-terminal mutations (Chapter 6).

CHAPTER 6

ROLE OF C-TERMINAL DOMAINS IN SURFACE ATTACHMENT OF THE FRUCTOSYLTRANSFERASE OF *STREPTOCOCCUS SALIVARIUS*¹

6.1 INTRODUCTION

Initial deletion studies implicated the C-terminal region of the Ftf in surface-attachment (Chapter 4). This region displays high homology with the C-termini of other Gram-positive surface-bound polypeptides such as the M-protein of *S. pyogenes* (Hollingshead *et al.*, 1986; Mouw *et al.*, 1988). In the case of M-protein, the C-terminus includes a cell wall sorting signal which consists of a consensus pentapeptide motif (LXPTG) followed by a C-terminal hydrophobic domain ending with a positively charged tail (Fig 6.1; Fischetti *et al.*, 1991). It is envisaged that the secretion of cell-surface proteins through the cytoplasmic membrane is hindered by the presence of the charged tail which allows the LPXTG consensus to be maintained in a position where it is proteolytically cleaved. Cleavage between the threonine and glycine results in cross-bridge formation between the threonine and pentaglycine of the peptidoglycan (Schneewind *et al.*, 1995; Pancholi and Fischetti, 1988; Mazmanian *et al.*, 1999). Such a consensus pentapeptide is absent in the Ftf of *S. salivarius* raising the question as to how the enzyme remains attached to the cell surface.

Besides the possible involvement of the hydrophobic domain in retention, another region common to surface-bound proteins lies directly N-terminal to the sorting signal. This so called wall-associated domain is considered to vary in length between 50-125 amino acids and is characterized by a high proline-glycine and threonine-serine ratio (Fischetti *et al.*, 1991). A similar extended domain spanning 178 amino acids is apparent at the C-terminus of the Ftf of *S. salivarius* (Fig 4.4; Fig 6.1). This domain is separated from the catalytic domain of the Ftf (defined as that region between Q¹⁹⁸ and D⁶³¹ possessing high homology with the SacBs of bacilli (Fig 4.4) by a spacer region of 73 amino acids.

In order to distinguish the roles of the hydrophobic and wall-associated domain in the attachment of the Ftf of *S. salivarius* to the surface of the cell, regions of the *ftf* gene coding for these domains were deleted and the recombinant proteins expressed in the heterologous host

¹ Some of the results of this chapter have already been published – Rathsam and Jacques, *J. Bacteriol* **180**: 6400-6403, 1998.

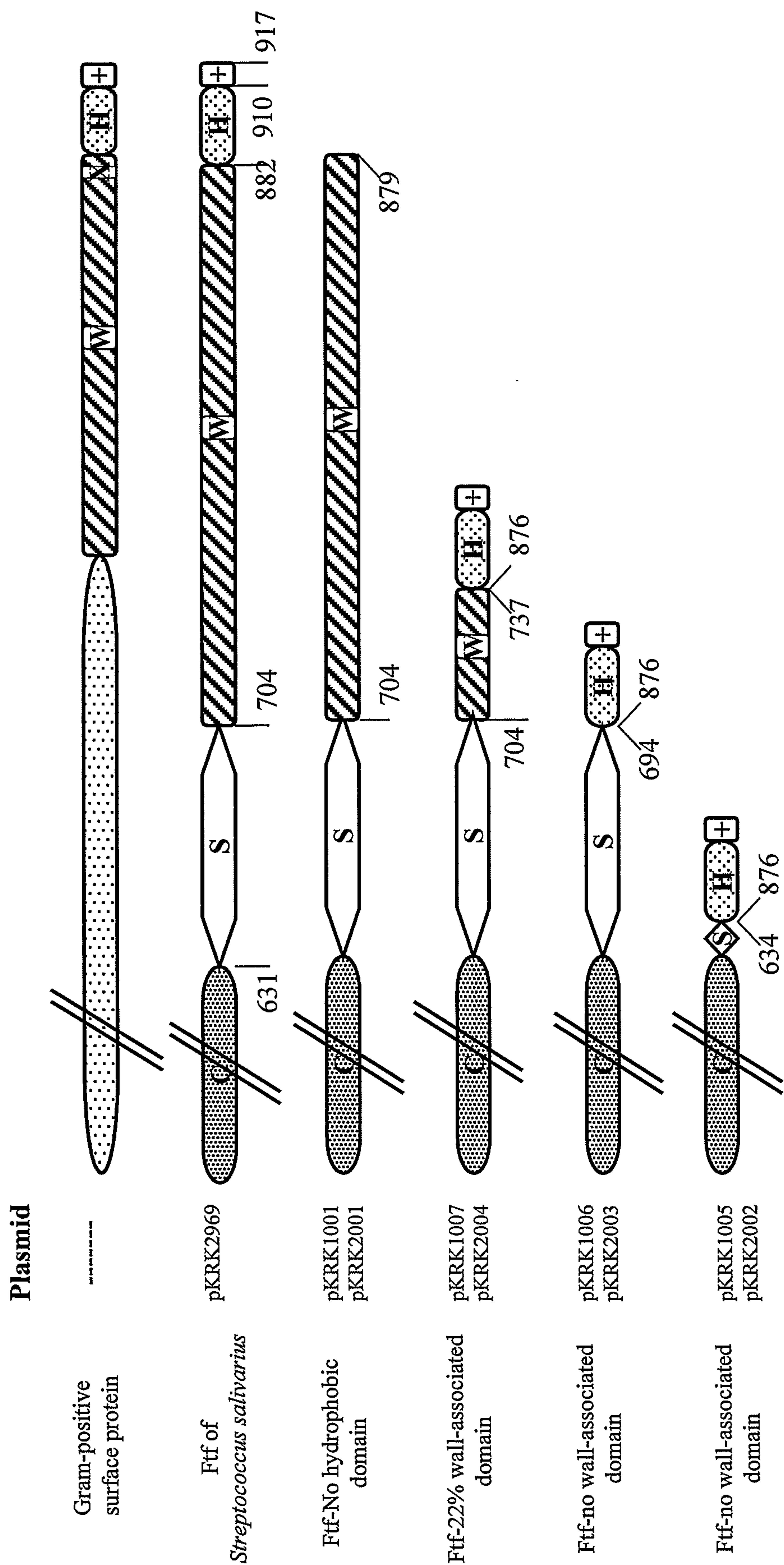


Figure 6.1 Comparison of the generalised domain structure of a Gram-positive wall-associated protein with that of the parental and mutated Ftfs of *S. salivarius*. C, catalytic domain; S, 72 amino acid spacer region; W, wall-associated domain; X, pentapeptide wall attachment motif (LXPTG); H, hydrophobic domain; +, positively charged amino acid C-terminus.

Streptococcus gordonii.

6.2 METHODS

6.2.1 Bacterial Strains and Growth Conditions

Escherichia coli NM522, *S. salivarius* ATCC 25975 and *S. gordonii* LGR2 (Table 2.1) were grown in liquid and on solid medium as previously described (2.3)

6.2.2 Plasmids and phagemids

Plasmids and phagemids (Table 2.1) were maintained in *E. coli* NM522. Cells harbouring pIBI31 or derivatives of the pKRK1000 series were grown with 100 µg ampicillin ml⁻¹, whilst cells harbouring pVA838 or derivatives of the pKRK2000 series were grown with 500 µg erythromycin ml⁻¹ and/or 25 µg chloramphenicol ml⁻¹. Transformation of *E. coli* by electroporation and natural transformation of *S. gordonii* LGR2 were achieved as previously described (2.4.14; 2.4.15).

6.2.3 Site-Directed Mutagenesis

Site-directed mutations of the *ftf* gene were obtained by the T7 modification of the TransformerTM Site-Directed Mutagenesis Kit supplied by Clontech Laboratories as previously described (5.2.3; Fig 5.1; Table 5.2).

6.2.4 Formation of Truncated *ftf* Genes

The introduction of a 'stop' codon in the *ftf* gene allowed the construction of a *ftf* allele that expressed a truncated Ftf which was devoid of its hydrophobic domain and its positively charged C-terminus (Fig 5.1; Fig 6.1), while the introduction of *Bam*HI sites into the *ftf* gene allowed in-frame deletions to be constructed in the C-terminus. These in-frame deletions expressed Ftfs devoid of portions of the proline-glycine / threonine-serine-rich wall-associated domain whilst maintaining their hydrophobic tails and positively charged C-termini in tact (Fig 5.1; Fig 6.1). Ftf⁺ transformants carrying the various *ftf* alleles were selected on BHI medium supplemented with 40 µl erythromycin ml⁻¹ and 3% (w/v) raffinose which allowed the detection of extracellular fructan in all but one case (see Section 6.3.1 and 6.3.2 below). Automated DNA sequencing was carried out on double-stranded DNA templates at the University of Sydney and Prince Alfred Molecular Analysis Centre (SUPAMAC) to validate the site-directed mutations and/or the nature of the in frame deletions. This was done prior to and following the subcloning of the altered *ftf*

genes into the *E. coli*-*Streptococcus* shuttle vector pVA838 and transforming into *S. gordonii*.

6.2.5 Detection of Ftf Activity

Ftf activity was detected either qualitatively or quantitatively as previously described (2.5.1; 2.5.2; 2.5.3)

6.2.6 Release of Cell-bound Ftf Activity in Washed Cell Suspensions

Washed suspensions of streptococcal cells grown in THB and sampled throughout growth were used to determine the release of cell-bound Ftf activity. Washed cell suspensions were incubated at 37°C with 10 mM sucrose for 5 min as described previously (Milward and Jacques, 1990).

6.2.7 Fractionation of *E. coli* Cells

Periplasmic extracts of *E. coli* were prepared by an adaptation of the method of Bedouelle and Duplay (1988). Essentially, 12 ml of cells from a 25 ml culture were harvested by centrifugation (10 000 g, 4°C, 10 min) and washed with 5 ml of 10 mM Tris HCl buffer pH 8.0 before being resuspended in 500 µl of 30 mM Tris HCl buffer pH 8.0 containing 20% (w/v) melezitose. EDTA was then added to a concentration of 1 mM and the suspension gently stirred for 10 min. The cells were then harvested by centrifugation (10 000 g, 4°C, 10 min) and osmotically shocked by re-suspending them and gently stirring for 10 min in 500 µl of 500 µM MgCl₂. The shocked cells were subsequently harvested by centrifugation (10 000 g, 4°C, 10 min) and the supernatant retained as the source of shock fluid or periplasmic extract. The cell pellet was washed with 500 µl of 30 mM Tris HCl buffer pH 8.0 containing 20% (w/v) melezitose and resuspended in 1-2 ml of 30 mM Tris HCl buffer pH 8.0 containing 20% (w/v) melezitose. The shocked cell suspensions were subsequently permeabilized by the addition of 100 µl chloroform and 50 µl of 0.1% (w/v) SDS per ml of resuspended cells followed by vortexing for 10 s. All supernatants and resuspended cells were kept on ice prior to assaying for Ftf activity. Permeabilized unshocked cells were similarly obtained and used as controls. Prior to assaying permeabilized cells for Ftf activity, Triton X-100 was added to a final concentration of 0.5% (w/v) of the actual assay reaction, in order to reverse the effects of the added SDS.

6.2.8 Data Base Searches

All DNA and protein comparisons made use of the programs available through ANGIS

(2.6).

6.3 RESULTS

6.3.1 Expression of Mutated Ftf in *E. coli*

The wall-associated domain of Gram-positive surface proteins is defined as a region spanning 50-125 residues with the content of proline-glycine and threonine-serine residues ranging between 15-32% and 13-38% respectively (Fischetti *et al.*, 1991). The wall-associated domain of the Ftf of *S. salivarius* that directly precedes the hydrophobic C-terminal domain contains an extended proline-glycine- and threonine-serine-rich domain spanning 178 amino acids from Pro-705 to Ser-882 (Fig 6.1). This domain contains a proline-glycine and threonine-serine content of 17% and 22% respectively. Mutated Ftf containing deletions in both the hydrophobic and wall-associated domains were expressed in *E. coli* (Fig 6.1). All mutated forms of the enzyme retained the ability to hydrolyze sucrose and to form fructan except the mutated Ftf expressed by pKRK1005 or pKRK2002 (Section 6.3.2 and Fig 6.2 below). This inactive form of the enzyme had lost the entire wall-associated domain as well as 96% of the C-terminal spacer region linking it to the catalytic domain. It retained its C-terminal hydrophobic domain (Fig 6.1).

Fractionation of *E. coli* cells indicated the importance of the hydrophobic domain in preventing Ftf export into the periplasm since deletion of this domain resulted in an increase in the percentage of the Ftf detected in the periplasmic fraction (Table 6.1). The result was consistent with the view that the C-terminal hydrophobic region with its positively charged tail acted to anchor the enzyme on the periplasmic side of the plasma membrane of *E. coli*. Extensive or total deletion of the wall-associated region reduced the amount of Ftf exported compared with the parental enzyme (Table 6.1). This result could be explained by a reduction in size of the Ftf protein as the mutated enzymes were smaller by 20% and 15% respectively. The results were consistent with the hypothesis that a significant reduction in size of a high M_r protein possessing a membrane-spanning hydrophobic C-terminus will positively enhance its retention on the periplasmic side of the plasma membrane after translocation (von Heijne, 1986).

6.3.2 Cellular Location of Mutated Ftf in *S. gordonii*

To study the effect of the deletions of various C-terminal domains on the localization of the protein in a *Streptococcus*, the mutated *ftf* genes were cloned into the *E. coli*-*Streptococcus* shuttle vector pVA838 and transformed into *S. gordonii*. Cells harvested at various stages throughout the growth cycle were used to determine the amount of Ftf bound to the cell as well

Table 6.1 Cellular location of Ftf activity expressed in *E. coli*^a

Plasmid	Ftf C-terminal Modification	Ftf Activity (%) ^b	
		Periplasm	Plasma Membrane
None	--	0	0
pKRK1969	None	42	58
pKRK1001	No hydrophobic domain	72	28
pKRK1006	No wall-associated domain	17	83
pKRK1007	22% wall-associated domain	14	86

^a Ftf activity for *E. coli* resuspended to an $A_{600} = 1.000$ and transformed with pKRK1969, pKRK1001, pKRK1006 or pKRK1007, was 972 - 1407 mU ml⁻¹, 1930 - 1126 mU ml⁻¹, 594 - 705 mU ml⁻¹ or 660 - 953 mU ml⁻¹ respectively.

^b Each assay was done in triplicate. For pKRK1969 and pKRK1001 the mean of triplicate experiments are shown, whilst for pKRK1006 and pKRK1007, the mean of duplicate experiments are shown.

as the amount secreted into the culture medium. The percentage of enzyme activity released from the surface of the cell in the presence of sucrose was also measured.

Deletion of the C-terminal hydrophobic domain resulted in the secretion of the Ftf by the heterologous host, *S. gordonii* (Fig 6.2). This result was in keeping with those obtained previously with C-terminal-deletion mutations constructed by random exonuclease III digestion of the *ftf* gene (4.3.4.3; plasmids pKRK2914 and pKRK2916 in Fig 6.2). In these instances the hydrophobic domain together with C-terminal portions of the wall-associated domain were removed. However, *S. gordonii* carrying pKRK2003 that encoded an Ftf devoid of its wall-associated domain but retaining its hydrophobic C-terminus in tact was also secreted (Fig 6.2). This indicated that the wall-associated domain was essential for the stable binding of the Ftf to the surface of *S. gordonii* irrespective of the presence of the hydrophobic region. Inclusion of 22% of the wall-associated domain together with the C-terminal hydrophobic domain resulted in 51% of the Ftf remaining attached to the surface of *S. gordonii* (Fig 6.2). This result further

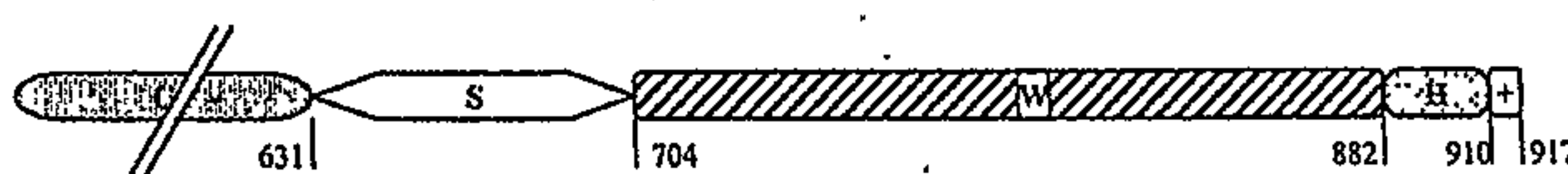
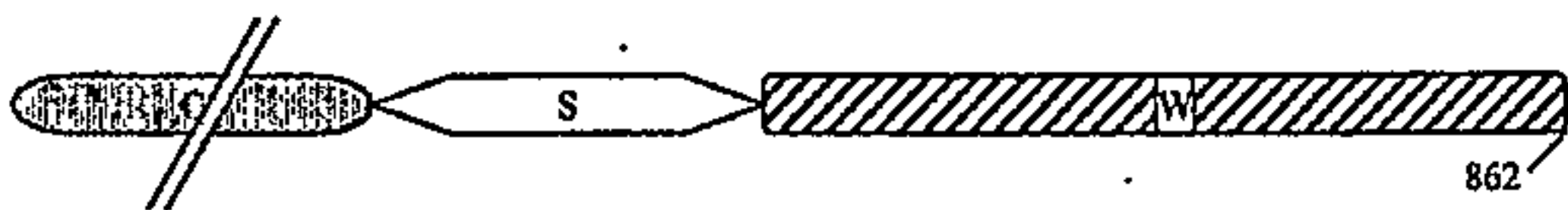
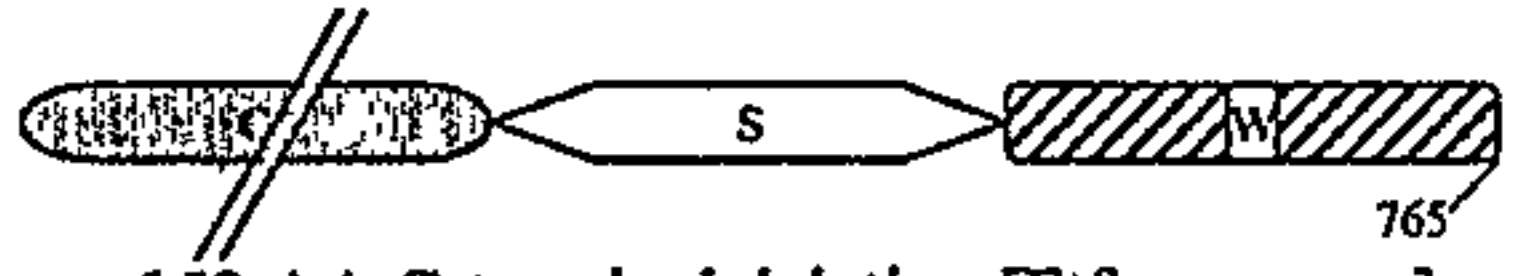
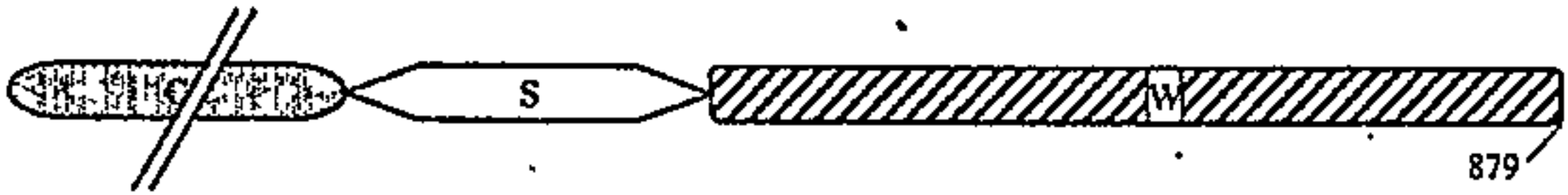
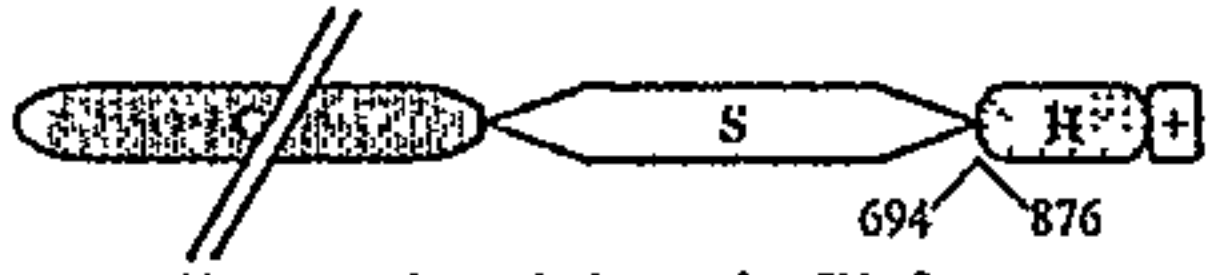
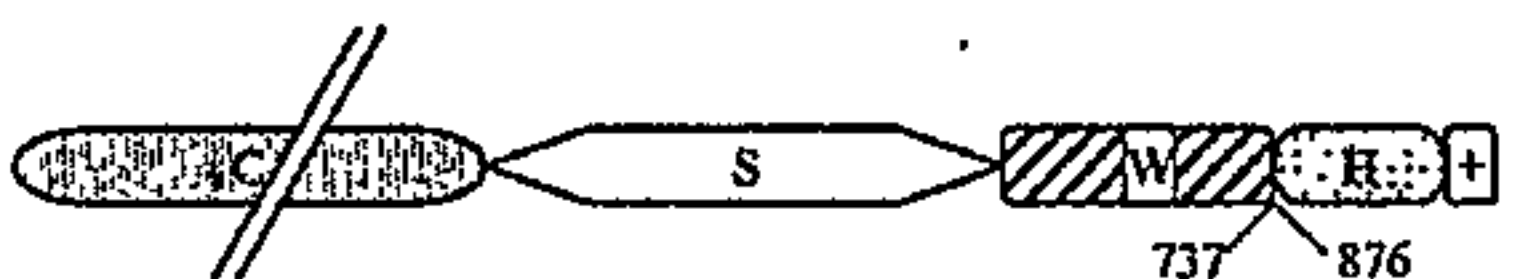
Bacterium	Plasmid	Domain Structure of Expressed Ftf	Extracellular Ftf Activity (%)	
			Cell Culture	Sucrose Incubation
<i>S. salivarius</i> <i>S. gordonii</i>	None pKRK2969	 Parental Ftf Parental Ftf	15±4 (3) 25±4 (3)	81±5 (3) 51±1 (3)
<i>S. gordonii</i>	pKRK2914	 55 AA C-terminal deletion [Ftf _(Δ863-917)]	89±4 (4)	---
<i>S. gordonii</i>	pKRK2816	 152 AA C-terminal deletion [Ftf _(Δ766-917)]	77±5 (4)	---
<i>S. gordonii</i>	pKRK2001	 No hydrophobic domain [Ftf _(Δ880-917)]	82±3 (3)	---
<i>S. gordonii</i>	pKRK2002	 No wall-associated or spacer domain [Ftf _(Δ635-875)]	No Activity (3)	No Activity (3)
<i>S. gordonii</i>	pKRK2003	 No wall-associated domain [Ftf _(Δ695-875)]	95±4 (3)	---
<i>S. gordonii</i>	pKRK2004	 22% wall-associated domain [Ftf _(Δ738-875)]	46±3 (3)	76±4 (3)

Figure 6.2 Comparison of the generalized domain structure of parental and mutated Ftfs and their cellular location prior to and following incubation of washed cells with sucrose. C, catalytic domain; S, 73 amino acid spacer; W, wall-associated domain; H, hydrophobic domain; +, positively charged amino acid C-terminus. The numbers refer to the first or last amino acid present within a given or truncated domain.

^a Ftf activity (mean ± SD for *n* repeat experiments) Ftf activity for *S. salivarius* was within the range of 580-660 mU (mg dry wt)⁻¹ and for *S. gordonii*, transformed with the pKRK2969, pKRK2914, pKRK2816, pKRK2001, pKRK2003 or pKRK2004 in the range 1220-1720 mU (mg dry wt)⁻¹, 700-1360 mU (mg dry wt)⁻¹, 610-920 mU (mg dry wt)⁻¹, 891-1095 mU (mg dry wt)⁻¹, 644-904 mU (mg dry wt)⁻¹ or 756-958 mU (mg dry wt)⁻¹ respectively.

^b At late exponential phase, *S. salivarius* and *S. gordonii* were harvested and the amount of Ftf activity in the supernatant compared with that bound to the cell.

^c Washed cells were incubated with 10 mM sucrose and the amount of Ftf released from the cell determined.

^d Data only shown where the majority of Ftf activity was initially cell-associated.

supported the hypothesis that the wall-associated domain was as important as the hydrophobic domain in stabilizing the Ftf on the surface of the cell. The cell-associated Ftf possessing only 22% of the wall-associated domain was released from the surface of *S. gordonii* by its substrate, sucrose, to a similar extent as the intact parental enzyme (Fig 6.2).

6.4 DISCUSSION

The absence of the C-terminal pentapeptide anchoring signal (LPXTG) found in other Gram-positive surface proteins implicated the importance of the wall-spanning region and the hydrophobic domain in attaching the Ftf of *S. salivarius* to the surface of the cell (Chapter 4). The results of this study confirmed the well documented role of the hydrophobic domain in anchoring proteins to the plasma membranes of both Gram-negative and Gram-positive bacteria (MacIntyre *et al.*, 1988; Hansson *et al.*, 1992). The partitioning experiments performed in *E. coli* showed that removal of the hydrophobic domain of the Ftf of *S. salivarius* resulted in export of the enzyme into the periplasm. The same mutated Ftf was also secreted by *S. gordonii*, confirming the importance of the hydrophobic domain in locating the Ftf on the surface of streptococci.

In contrast to the role of the LPXTG pentapeptide and the hydrophobic domain in stabilizing proteins on the surface of Gram-positive bacteria, the role of the proline-glycine-, threonine-serine-rich wall-associated domain has remained speculative. It has been suggested that this domain might also augment the stability of the cell-bound form of the enzyme by allowing it to intercalate with the peptidoglycan matrix (Burne and Penders, 1992). Removal of this domain from the Ftf of *S. salivarius* along with 86% of the 72 amino acid C-terminal spacer region resulted in an inactive enzyme. The association of the hydrophobic domain in direct juxtaposition to the catalytic domain of the Ftf may have destabilized the tertiary structure of the enzyme. This hypothesis is supported by the observation that an altered Ftf truncated at Ser-609 (expressed by pKRK1661), and thus possessing a 21 amino acid C-terminal deletion of the 'catalytic' domain, was also found to be inactive (4.3.4.1). These two observations suggest that the tertiary structure formed by the C-terminal amino acids of the catalytic domain of the Ftf is critical for the maintenance of catalytic function.

In order to further investigate the role of the wall-associated domain in surface attachment, two other *ftf* genes were constructed that expressed Ftfs devoid of the wall-associated domain or possessing only 22% of it. Both of these proteins contained the intact hydrophobic C-terminal domain. These two Ftfs retained activity when expressed in either *E. coli* or *S. gordonii*. The

smaller of these proteins devoid of the entire cell-wall associated domain was secreted by *S. gordonii* (87%) supporting the hypothesis that the wall-associated domain was just as essential as the hydrophobic C-terminal domain in stabilizing surface attachment in the high M_r Ftf. This was confirmed by expression of the larger of the two mutated Ftfs possessing 22% of the wall-associated domain intact. In this instance, half of the Ftf was retained on the cell surface.

This result was similar to that observed for the role of the C-terminal domains of *spaP* which codes for the cell surface antigen P1 of *S. mutans*. The gene product of a *spaP* deletion mutant created by Homonylo-McGavin and Lee (1996) devoid of the C-terminal domains and expressed in P1-negative *S. mutans* and *S. gordonii*, was only found in the culture supernatant. A truncated P1 containing the 44 amino acid wall-spanning domain, but missing the C-terminally charged tail and hydrophobic domain (including the LPXTG anchor consensus), was 47% retained on the surface of *S. gordonii* (Homonylo-McGavin and Lee, 1996).

The presence of an extended wall-associated domain possessing turn-promoting proline-glycine residues could theoretically allow the C-terminal wall-associated region of the Ftf to span the cell wall by randomly intercalating throughout the peptidoglycan-carbohydrate-teichoic acid matrix. The high serine-threonine content could further stabilize this association by allowing hydrogen bonding to the carbohydrate constituents of the cell wall. Whether this is the case or not, it is clear from the results of this study that both the C-terminal hydrophobic domain and the extended wall-associated domain of the Ftf of *S. salivarius* are required for stable attachment of the enzyme to a Gram-positive cell surface.