

TARGET SEQUENCE REQUIREMENTS FOR HIGH FREQUENCY, SITE-
SPECIFIC TRANSPOSITION OF Tn7

By

Mohammad Ishtiaq Qadri

Centre for Biotechnology
Imperial College of Science and Technology
London SW7 2AZ

January, 1988

A dissertation submitted to the University of London in
candidature for the degree of Doctor of Philosophy and to
the Imperial College of Science and Technology in
candidature for the Diploma of Imperial College.

ABSTRACT

The aim of the research was to define the target sequence requirements for high frequency site-specific transposition of the 14 Kb bacterial transposon Tn7. This transposon confers resistance to trimethoprim, spectinomycin and streptomycin. Previous studies have suggested that Tn7 inserts randomly into plasmid DNA with a low frequency and into a unique site in the chromosome of Escherichia coli with high frequency. Site-specific transposition of Tn7 has also been shown in the chromosomes of Pseudomonas aeruginosa, Caulobacter crescentus and Xanthomonas campestris pv campestris.

In this study a variety of approaches were taken to define the target sequence requirements for high frequency Tn7 transposition to the attachment site (attTn7). A novel M13 based assay system was developed to measure the transposition frequency of Tn7. Use of this assay permitted a 64 bp region within the E. coli attTn7 to be defined which serves as an efficient and site-specific target for Tn7 transposition.

Tn7 insertional properties were analysed in three bacterial species i.e. Agrobacterium tumefaciens, Klebsiella pneumoniae and Serratia marcescens. Tn7 insertion was found to occur at a unique site in the chromosomes of A. tumefaciens and K. pneumoniae while two specific sites were

found in the genome of S. marcescens. All of these attTn7 sites were cloned in the broad host range plasmid, pTJS75Kn.

DNA sequence analysis of attTn7 sites from K. pneumoniae and S. marcescens (1st site) show that Tn7 insertion occurs in the 3' end of a sequence which by homology to the E. coli site probably encodes the C-terminus of glucosamine synthetase (glmS). The second S. marcescens attTn7 site has no such homology.

Comparison of attTn7 DNA sequence of K. pneumoniae and S. marcescens (1st site) with the 64 bp region of E. coli attTn7 showed that the point of insertion of Tn7 is an integral number of helical turns away from one highly conserved domain (in the glmS box). In order to test the requirement for an integral number of helical turns between the glmS box and the insertion site, a ClaI restriction site was generated in the intervening sequence of E. coli attTn7 to provide a cloning site and so allowed the separation of two domains with altered spacings. These spacing mutants however retained wild-type attTn7 properties with respect to the frequency of transposition.

However, a requirement for a second, structure-specific domain (an inverted repeat) was identified within the attTn7 site region. This structure appears to be required for full attTn7 site activity.

Individual pairs of nucleotides responsible for high

frequency transposition of Tn7 were defined within a 70 bp mini attTn7 site, by random, region specific mutagenesis. A large number of point and multiple mutants were tested using the M13 transduction assay and three regions, which may be required for full activity, were identified within the 36 bp glmS box of the attTn7. The identified nucleotides probably participate in a highly precise DNA-protein complex for site-specific transposition of Tn7.

ACKNOWLEDGEMENTS

First of all, I wish to express my gratitude to my supervisor Dr. Conrad Lichtenstein for his valuable suggestions, guidance, constructive criticism and encouragement throughout the course of this study. Completion of this work would have been difficult without his personal interest and affection which has been shown to me since the very beginning.

I would also like to thank Prof. Brian Hartley for providing me the opportunity to work in the prestigious Centre for Biotechnology and allowing the use of facilities provided for his group.

I gratefully acknowledge the intellectual contributions of Carlos Flores, Tony Wright and Tony Day which was made for this thesis. Their suggestions, sincere help and friendship were always available to me. Thanks are due to Terese Anderton for showing me the DNA sequencing techniques.

Special thanks to my friends Hasan Zaidi, Elizabeth Black, Sherry Fuller, Jane Clarke-Day, Jeremy Green, Samuel Amarthi, Sanjoy Ray and Tasneem Gohir for their help, sense of humour and for tolerating me. My four years at Imperial College have been most enjoyable and I acknowledge all those who shared joyful moments with me.

I would also like to express my thanks to Glynn Millhouse for his excellent photography skills and Kevin Cope for his prompt ordering of reagents needed for this project.

And last, but not the least, Alison Davis for everything.

TABLE OF CONTENTS

TITLE PAGE	1
ABSTRACT	2
ACKNOWLEDGEMENTS	5
TABLE OF CONTENTS	6
LIST OF FIGURES	15
LIST OF TABLES	21
ABBREVIATIONS	23
CHAPTER 1: INTRODUCTION	
1.1 Definition, History and Nomenclature.....	27
of Transposable Elements.	
1.2 Consequences of Transposition.....	33
1.2.1 DNA Rearrangements Associated.....	33
with Transposable Elements.	
1.2.2 Alteration of Gene Expression.....	38
by Transposable Elements.	
1.3 Structural Organisation and Functions.....	39
of Prokaryotic Transposable Elements.	
1.3.1 Transpositional Immunity.....	40
1.3.2 Target Sequence Duplication at.....	41
the Sites of Insertion of	
Transposable Elements.	
1.3.3 Transposable Element Encoded.....	43
Transposition Proteins.	
1.4 Specificity for Insertion of.....	46
Transposable Elements.	
1.4.1 Random Insertion of Transposable....	48
Elements.	
1.4.2 Region Specific Insertion of.....	52
Transposable Elements.	
1.4.3 High Specificity Insertion of.....	54
Transposable Elements.	
1.5 Control of Transposition.....	59

1.6	Models for Transposition.....	63
1.6.1	Replicative Mode of Transposition.....	64
1.6.2	Conservative Mode of Transposition....	68
1.7	Site-specific Recombination.....	71
1.7.1	Site-specific Integration of..... Bacteriophages into the Genome of Bacteria.	72
1.7.2	Cointegrate Resolution; Another..... Site-specific Recombination System.	75
1.7.3	Site-specific Inversion of DNA.....	78
1.7.4	Requirements for Site-specific..... Recombination.	79
1.7.5	DNA Flexibility as a Determinant..... in Site-specific Recombination.	83
1.8	Eukaryotic Transposable Elements.....	85
1.9	Evolution of Prokaryotic Transposable.... Elements.	88
1.10	Tn7 Transposition and Requirements.....	91
1.11	Preliminary Account of the Project.....	102

CHAPTER 2: MATERIALS AND METHODS

2.1	Materials.....	104
2.1.1	Consumables.....	104
2.1.2	Media.....	105
2.1.3	List of Bacterial Strains.....	107
2.1.4	List of Plasmids.....	109
2.1.5	M13 Derived Clones.....	112
2.2	Methods.....	117
2.2.1	Microbiological Techniques.....	117
2.2.1(i)	Antibiotic Selection.....	117
2.2.1(ii)	Transformation of <u>E.coli</u> ... with Plasmid DNA.	117
2.2.1(iii)	Assay for the Loss of..... Colicin E1 Production.	119
2.2.1(iv)	Bacterial Conjugation.....	120
2.2.1(v)	Curing of Episome and..... Plasmid.	120
2.2.1(vi)	M13 Transduction..... Frequency Assay of Tn7.	121
2.2.2	Biochemical Techniques.....	123

2.2.2A	General Techniques.....	123
2.2.2A(i)	Extraction of DNA with.....	123
	Phenol and Phenol/ Chloroform Mixture.	
2.2.2A(ii)	Precipitation of DNA.....	123
	with Ethanol.	
2.2.2A(iii)	Trichloroacetic Acid.....	124
	(TCA) Precipitation.	
2.2.2A(iv)	Scintillation Counting.....	125
2.2.2A(v)	Spectrophotometry.....	125
2.2.2A(vi)	Autoradiography.....	126
2.2.2B	DNA Preparation.....	126
2.2.2B(i)	Preparation of Covalently.....	126
	Closed Supercoiled Plasmid DNA.	
2.2.2B(ii)	Quick Plasmid DNA.....	128
	Preparation.	
2.2.2B(iii)	Preparation of.....	129
	Chromosomal DNA from <u>E. coli</u> and Other Bacterial Species.	
2.2.2B(iv)	Isolation of M13 Single-..	131
	stranded DNA.	
2.2.2B(v)	Isolation of M13.....	132
	Replicative Form DNA.	
2.2.2B(vi)	Purification of DNA.....	132
	Restriction Fragments from Agarose Gels.	
2.2.2B(vii)	Purification of DNA.....	134
	Restriction Fragments from Acrylamide Gels.	
2.2.2C	Gel Electrophoresis.....	134
2.2.2C(i)	Useful Molecular Markers..	135
2.2.2C(ii)	DNA Electrophoresis on...	135
	Horizontal Agarose Gels.	
2.2.2C(iii)	Polyacrylamide Gels.....	136
2.2.2C(iv)	Non-Denaturing Thin.....	136
	Acrylamide Gels.	
2.2.2C(v)	Denaturing Gradient.....	137
	Acrylamide Gels.	
2.2.2C(vi)	Non-Denaturing Gradient...	139
	Thick Acrylamide Gels.	
2.2.2C(vii)	Denaturing Thin Acryla-...	139
	amide Gels for DNA Sequencing.	
2.2.2C(viii)	Analysis of Proteins on...	140
	12.5% Discontinuous SDS Polyacrylamide Gel.	

2.2.2D	Modification of DNA by Enzyme.....	141
	Catalysis.	
2.2.2D(i)	Restriction Endonuclease..	141
	Digestion of DNA.	
2.2.2D(ii)	<u>In Vitro</u> Ligation of DNA..	142
	Using T4 DNA Ligase.	
2.2.2D(iii)	Incorporation of Radio-....	142
	active Nucleotide into	
	DNA Nick Translation.	
2.2.2D(iv)	Incorporation of [Gamma...]	143
	³² P] ATP at the 5' End of	
	Oligonucleotides Using	
	T4 Polynucleotide Kinase.	
2.2.2D(v)	Dephosphorylation of the..	143
	5' Terminus of DNA.	
2.2.2D(vi)	Preparation of Labelled..	144
	DNA Restriction Fragments	
	Using DNA Polymerase 1,	
	Klenow Fragment.	
2.2.2D(vii)	Unidirectional Deletion...	144
	Analysis Using	
	Exonuclease III and	
	S1 Nuclease.	
2.2.2D(viii)	DNA Sequencing.....	145
	(a) DNA Sequencing.....	146
	Reaction Mixture.	
	(b) Annealing and.....	146
	Polymerisation.	
	(c) Gel Loading.....	147
2.2.2E	Cloning Into M13 Phage Vectors.....	148
2.2.2E(i)	Identification of the.....	149
	Insert in M13	
	Phage Vectors.	
2.2.2E(ii)	Identification of M13.....	149
	Clones with Insert	
	in Opposite Orientation.	
2.2.2F	<u>In Vitro</u> Hybridisation Techniques.....	150
2.2.2F(i)	Southern Blot Analysis.	151
2.2.2F(ii)	M13 Plaque Lift.....	152
	Hybridisation.	
2.2.2F(iii)	<u>In Situ</u> Colony.....	153
	Hybridisation.	
2.2.3	Base Substitution Mutagenesis Techniques.....	153
2.2.3A	Site-Specific, Mismatched.....	153
	Oligonucleotide Directed	
	Mutagenesis.	
2.2.3A(i)	Purification of the.....	154

	Heptadecadmer.	
2.2.3A(ii)	Preparation of Gel and.....	154
	Electrophoresis.	
2.2.3A(iii)	Estimation of.....	155
	Oligonucleotide	
	Concentration.	
2.2.3A(iv)	Use of Oligonucleotide.....	156
	as a Mutagenic Primer.	
2.2.3B	Region Directed Random Mutagenesis.....	157
	Techniques.	
2.2.3B(i)	Forced Misincorporation...	157
2.2.3B(ii)	Myers Mutagenesis.....	158
	Technique.	
2.2.3B(iii)	Use of Sodium Bisulphite..	160
	as a Mutagen to Introduce	
	Base Substitution.	
	(a) Preparation of the.....	160
	Gapped Circular	
	Molecules for Target	
	Mutagenesis.	
	(b) Procedure for Sodium..	161
	Bisulphite Treatment.	
	(c) Identification of.....	162
	the Mutants.	
2.2.4	<u>In Vitro</u> DNA-Protein Interaction.....	163
	Studies.	
2.2.4(i)	Preparation of Crude.....	163
	Extracts.	
2.2.4(ii)	Formation of DNA-Protein..	164
	Complex <u>In Vitro</u> .	

CHAPTER 3: THE USE OF SINGLE-STRANDED FILAMENTOUS M13 PHAGE AS TRANSDUCING PARTICLES FOR DETER- MINING THE FREQUENCY OF TRANSPOSITION OF Tn7 IN VIVO

3.1	Introduction.....	165
3.2	Advantages of Using M13 Phages.....	166
3.3	Factors Affecting the Transposition of Tn7..	169
	into M13 Phages.	
3.3.1	Excision and Deletion Associated...	170
	with Transposition of Tn7 in M13	
	Phages.	
3.3.2	Time Course for Transposition of.....	172
	Tn7.	

3.4	Development of the Transduction Assay.....	174
	in M13 Phages Containing <u>E. coli</u> <u>attTn7</u> Site.	
3.5	M13 Transduction Assay on Junction <u>attTn7</u> ... Sites (Tn7- <u>attTn7</u> Fragment).	179
3.6	M13 Transduction Assay of Tn7 Utilising..... Different Bacterial Host Strain.	186

CHAPTER 4: DEFINITION OF THE MINIMAL TARGET DNA SEQUENCE FOR HIGH FREQUENCY TRANSPOSITION OF Tn7 TO THE E.COLI ATTn7 SITE

4.1	Introduction.....	190
4.2	Deletion Mutagenesis.....	191
4.3	The Minimum <u>attTn7</u> is in a 64 bp Fragment...	194
4.4	Cloning of a 70 bp <u>attTn7</u> Site into M13..... Phage Vector, mp 10.	207

CHAPTER 5: ANALYSIS OF Tn7 TRANSPOSITION INTO THE CHROMOSOME OF AGROBACTERIUM TUMEFACIENS, KLEBSIELLA PNEUMONIAE, SERRATIA MARCESCENS AND TO SECONDARY SITES IN ESCHERICHIA COLI

5.1	Introduction.....	211
5.2	Experimental Strategy for the Selection.... for Tn7 Transposition to the Bacterial Chromosome.	213
5.2.1	Step-1; Conjugal Transfer of..... RP4::Tn7.	215
5.2.2	Step-2; Exclusion of RP4::Tn7 by..... Conjugal Transfer of Incompatible Plasmid pPH1JI.	217
5.3	Measurement of Tn7 Transposition Frequency..	219
5.4	Determination of Tn7 Target Replicon.....	220
5.5	Southern Blot Analysis to Determine Site-... specificity of Tn7 Transposition to the Chromosome.	222
5.6	Tn7 has Two Insertion Sites (<u>attTn7</u> sites)..	235

in S. marcescens.

5.7 Tn7 has a Secondary <u>attTn7</u> Site in the...	39
Chromosome of <u>E. coli</u> .	
5.8 Experimental Conclusions.....	243

CHAPTER 6: DNA SEQUENCE DETERMINATION OF THE Tn7
INSERTION SITES IN THE CHROMOSOMES OF
A.TUMEFACIENS, K.PNEUMONIAE AND
S.MARCESCENS

6.1 Introduction.....	245
6.2 Strategy for Cloning the Insertion Sites....	246
of Tn7 from the Chromosomes of Three Bacterial Species.	
6.2.1 Cloning of the <u>attTn7</u> Site from.....	249
the Chromosome of <u>A. tumefaciens</u> .	
6.2.2 Cloning of the <u>attTn7</u> Sites from.....	250
the Chromosome of <u>K. pneumoniae</u> .	
6.2.3 Cloning of the <u>attTn7</u> Sites from.....	251
the Chromosome of <u>S. marcescens</u> .	
6.3 DNA Sequence Analysis of <u>attTn7</u> from the....	255
Chromosome of the Three Bacterial Strains.	
6.3.1 The Construction of <u>Sau3AI</u> Library...255	
of pAIQ1Δ::Tn7; pKIQ1::Tn7- pKIQ6::Tn7 and pSIQI::Tn7/ pSIQII::Tn7 in M13, mp10.	
6.3.2 Identification of Right End.....	256
Junction Fragments from the <u>Sau3AI</u> Library of pAIQ1Δ::Tn7; pKIQ1::Tn7- pKIQ6::Tn7 and pSIQI::Tn7/ pSIQII::Tn7.	
6.3.3 Identification of the Left End.....	258
Junction Fragment from the <u>Sau3AI</u> Library of pAIQ1Δ::Tn7; pKIQ1::Tn7- pKIQ6::Tn7 and pSIQI::Tn7/ pSIQII::Tn7.	
6.4 DNA Sequencing Analysis of the Right and...259	
Left Junction Fragments of Tn7- <u>attTn7</u> .	
6.5 Analysis of <u>attTn7</u> Sites.....	266

CHAPTER 7: EFFECT OF ALTERED INTERDOMAIN SEQUENCE
AND THE DELETION OF GLMS TRANSCRIPTIONAL
TERMINATOR IN THE ATTn7 ON THE
TRANSPOSITION OF Tn7

7.1	Introduction.....	272
7.2	Generation of <u>Clal</u> RE Site 5 bp..... Downstream of the 36 bp <u>glmS</u> Box in 70 bp <u>attTn7</u> Site (MAS).	274
7.3	Construction of Altered Spacing Clones..... Between the End of 36 bp <u>glmS</u> Box and the Point of Insertion of Tn7 in MASC.	276
7.4	Construction of Deletion Mutants in the..... <u>glmS</u> Transcriptional Terminator in 70 bp <u>attTn7</u> Site and Effect of These Deletions on Transduction Frequencies.	287
7.5	Where is Tn7 Inserted in Altered Spacing.... and Deletion Clones?	292
7.6	Conclusions from Altered Spacing and..... Deletion Experiments.	294

CHAPTER 8: EFFECT OF NUCLEOTIDE SUBSTITUTION
IN THE ATTn7 SEQUENCE ON TN7
TRANSPOSITION AND STUDIES ON DNA-PROTEIN
INTERACTIONS INVOLVING MINI ATTn7 SITE

8.1	Introduction.....	297
8.2	Experimental Approaches Adopted for..... <u>In Vitro</u> Region Specific Mutagenesis.	299
8.3	Use of Sodium Bisulphite as a Mutagen.....	308
8.3.1	Alteration of C to T in the Watson.... Strand of 70 bp <u>attTn7</u> DNA (MAS).	311
8.3.2	Alteration of C to T in the Crick.... Strand of the 89 bp <u>attTn7</u> DNA (Δ L849).	314
8.4	Distribution of Base-Pair Substitutions..... (C to T and G to A) in the 70 bp <u>attTn7</u> DNA.	318
8.5	Discussion on the Results of Mutagenesis... Studies.	319
8.6	<u>In Vitro</u> DNA-Protein Interaction Studies....	331

8.6.1 Experimental Aspects.....	331
CHAPTER 9: CONCLUDING REMARKS AND FUTURE EXPERIMENTS	
9.1 Summary of Results.....	335
9.2 Nature of the <u>att</u> Tn7 DNA for Target.....	346
Sequence Selection.	
9.3 Synapsis at the <u>att</u> Tn7 DNA and the Strand...348	
Cleavage at the Point of Insertion.	
9.4 Implication of DNA-Protein Interaction.....	355
9.5 Future Experiments.....	356
REFERENCES.....	361

LIST OF FIGURES

Chapter 1

1.1	Three Classes of Transposable Elements.	31
1.2	Types of Transposable Elements Promoted Events.	35
1.3	Intra-molecular Transposition by a "Cut and Paste" Mechanism.	37
1.4	Regulatory Mechanisms Used by Four Transposons to Control Their Transposition.	62
1.5	Replicative and Conservative Transposition.	64
1.6	Simplified Version of Symmetric and Asymmetric Models of Transposition.	66
1.7	Cointegrate Formation and Resolution of Tn3.	67
1.8	Model of Conservative Transposition of Tn5.	69
1.9	The <u>res</u> Site of Gammadelta.	76
1.10	Genetic Map of the <u>att</u> Tn7 Site from <u>E. coli</u> Chromosome.	96
1.11	Genetic Organisation of Transposon Tn7.	101

Chapter 2

2.1	Electroelution Tank.	133
-----	----------------------	-----

Chapter 3

3.1	Diagrammatic Representation of M13 Transduction Frequency Assay.	176
-----	--	-----

3.2	Graphical Representation of the.....	182
	Results of M13 Transduction Frequency Assay on Different <u>attTn7</u> Sites.	
3.3	Summary of the Results of M13.....	183
	Transduction. Frequency Assay of Different <u>attTn7</u> Sites.	

Chapter 4

4.1	Strategy of Unidirectional Deletion.....	195
	Analysis Using Exonuclease III and S1 Nuclease.	
4.2	DNA Sequencing Gels of Deletion.....	198
	Mutants.	
4.3	Graphical Representation of the.....	200
	Results of M13 Transduction Frequency Assays on Deletion Mutants of 958 bp <u>attTn7</u> Site.	
4.4	DNA Sequence Adjacent to Deletion.....	203
	Mutants $\Delta R11$ to $\Delta R38$.	
4.5	Summary of Unidirectional Deletion.....	205
	Analysis of 958 bp <u>attTn7</u> .	
4.6	Strategy for Cloning the 70 bp Mini.....	208
	<u>attTn7</u> Sequence from Deletion Mutant $\Delta L849$ in M13, mp10.	
4.7	DNA Sequence of the 70 bp Cloned....	210
	Mini <u>attTn7</u> site.	

Chapter 5

5.1	Phylogenetic Relationship Between.....	212
	the Bacterial Species Studied for Site-specific Transposition of Tn7.	
5.2	Diagrammatic Representation of.....	216
	Experimental Strategy for the Selection for Tn7 Transposition to the Bacterial Chromosome.	
5.3	Southern Blot of <u>EcoRI</u> RE Digested.....	225
	Genomic DNA of <u>A. tumefaciens</u> ::Tn7 Probed with Radiolabelled ColE1::Tn7 DNA.	

5.4	Southern Blot of <u>EcoRI</u> RE Digested.....	226
	Genomic DNA of <u>A. tumefaciens</u> ::Tn7 Probed with Radiolabelled pRS5 DNA.	
5.5	Southern Blot of <u>EcoRI</u> RE Digested.....	227
	Genomic DNA of <u>K. pneumoniae</u> ::Tn7 Probed with Radiolabelled ColE1::Tn7 DNA.	
5.6	Southern Blot of <u>EcoRI</u> RE Digested.....	228
	Genomic DNA of <u>K. pneumoniae</u> ::Tn7 Probed with Radiolabelled pRS5 DNA.	
5.7	Southern Blot of <u>EcoRI</u> RE Digested.....	229
	Genomic DNA of <u>S. marcescens</u> ::Tn7 Probed with Radiolabelled ColE1::Tn7 DNA.	
5.8	Southern Blot of <u>EcoRI</u> RE Digested.....	230
	Genomic DNA of <u>S. marcescens</u> ::Tn7 Probed with Radiolabelled <u>EcoRI</u> Left End Fragment (8.9 Kb) of Tn7 DNA.	
5.9	Southern Blot of <u>HindIII</u> , <u>BamHI</u> REs.....	232
	Digested Genomic DNAs of <u>A. tumefaciens</u> ::Tn7 (A::Tn7), <u>K. pneumoniae</u> ::Tn7 (K::Tn7) and <u>S. marcescens</u> ::Tn7 (S::Tn7) Probed with Radiolabelled ColE1::Tn7 DNA.	
5.10	Southern Blot of <u>SalI</u> , <u>SstII</u> , <u>SmaI</u>	233
	and <u>XhoI</u> REs Digested Genomic DNAs of <u>A. tumefaciens</u> ::Tn7 (A::Tn7) <u>K. pneumoniae</u> ::Tn7 (K::Tn7) and <u>S. marcescens</u> ::Tn7 (S::Tn7) Probed with Radiolabelled ColE1::Tn7 DNA.	
5.11	Southern Blot of <u>HindIII</u> , <u>SalI</u> and.....	236
	<u>SstII</u> REs Digested Genomic DNA of <u>S. marcescens</u> ::Tn7 Probed with Radiolabelled ColE1::Tn7 and pRS5 DNAs.	
5.12	Southern Blot of <u>EcoRI</u> , <u>HindIII</u> REs.....	242
	Digested Genomic DNA of $\Delta 7$::Tn7 Probed Separately with Radiolabelled ColE1::Tn7, pFK::Tn7 and pRS5 DNAs.	

Chapter 6

6.1	Restriction Maps of Plasmids pRK404,....	247
	pUC18 and pTJS75Kn.	

6.2	Photograph and Southern Blot of.....253
	<u>Sal</u> IRE Digested pTJS75Kan Derived Recombinant Plasmids (pSIQ1::Tn7- pSIQ12::Tn7) Probed with Radiolabelled RP4 DNA.
6.3	DNA Sequencing Gels of Tn7- <u>att</u> Tn7.....260
	(Left and Right) Junction Fragments from <u>A. tumefaciens</u> ::Tn7, <u>K. pneumoniae</u> ::Tn7 and <u>S. marcescens</u> ::Tn7 (both sites).
6.4	DNA Sequence of <u>att</u> Tn7 Sites of.....261
	<u>A. tumefaciens</u> and <u>K. pneumoniae</u> .
6.5	DNA Sequence of <u>att</u> Tn7 (1st site).....262
	of <u>S. marcescens</u> .
6.6	DNA Sequence of <u>att</u> Tn7 (2nd site).....263
	of <u>S. marcescens</u> .
6.7	Predicted Amino Acid Sequence at the....264
	C-terminal End of <u>glmS</u> Coding Region in the <u>att</u> Tn7 Sites of <u>E.coli</u> , <u>K. pneumoniae</u> and <u>S. marcescens</u> (1st site).
6.8	Comparison of Chromosomal <u>att</u> Tn7.....265
	Sites of <u>E.coli</u> (E.c.), <u>K.pneumoniae</u> (K.p.) and <u>S.marcescens</u> (S.m.) 1st site.

Chapter 7

7.1	Sequence of the Mutagenic Primer.....274
	Above the Mini <u>att</u> Tn7 Sequence.
7.2	Strategy and Results for Introducing....275
	the <u>Cla</u> I RE Site Inbetween the End of the <u>glmS</u> Coding Region and the Point of Insertion of Tn7.
7.3	Strategy for Cloning a 55 bp.....277
	Polylinker at the Newly Generated <u>Cla</u> I RE site in Mini <u>att</u> Tn7 DNA.
7.4	DNA Sequencing Gels of Altered.....280
	Spacing Clones (MASCPL77 and Derivatives).
7.5	DNA Sequencing Gel of Deletion.....281
	Mutants Δ1LOOP, Δ2LOOP and Δ3LOOP.

7.6	DNA Sequence of Altered Spacing.....	282
	Clones (MASCPL77 and Derivatives) and Three Deletion Clones (Δ 1LOOP, Δ 2LOOP and Δ 3LOOP).	
7.7	Graphical Representation of the.....	284
	Results of M13 Transduction Frequency Assays on Altered Spacing and Deletion Clones.	
7.8	Results of Linker Spacing Mutagenesis...	285
	of MAS.	
7.9	Potential Cruciform Structures Around... the 5 bp Target Duplication of Tn7 in <u>att</u> Tn7 Sites.	289

Chapter 8

8.1	Diagrammatic Representation of Forced... Misincorporation Mutagenesis.	300
8.2	A Perpendicular Gradient (10% to 80%)... Denaturant Acrylamide Gel.	305
8.3	Watson and Crick Strands of Target..... <u>att</u> Tn7 DNA for Mutagenesis Studies.	309
8.4	Strategy for Construction of M13..... Phage Vectors, MP11 <u>ATT</u> GC and MP11GC for Sodium Bisulphite Mutagenesis.	310
8.5	Strategy Adopted for Sodium..... Bisulphite Mutagenesis.	312
8.6	Autoradiograph of DNA Sequencing..... Gels Showing Mutants of the Watson Strand of <u>att</u> Tn7.	315
8.7	Autoradiograph of DNA Sequencing (i).. Gels Showing Mutants of the Crick (ii). Strand of <u>att</u> Tn7.	316 317
8.8	Results of Mutagenesis Studies on..... the Watson Strand of <u>att</u> Tn7 DNA.	320
8.9	Results of Mutagenesis Studies on..... the Crick Strand of <u>att</u> Tn7 DNA.	321
8.10	Important Guanosine and Cytidine..... Residues in <u>glm</u> S 36 bp Box.	328

- 8.11 Mutants in glmS Transcriptional.....330
Terminator; Their ΔG Values and
Correlation With Transposition
Frequency.
- 8.12 Binding Gel Showing 70 bp DNA Fragment..333
Containing the attTn7.

Chapter 9

- 9.1 DNA Sequence Comparison at Three....347
Domains within the 36 bp of glmS
Box of E. coli (E.c) attTn7 with
that of K. pneumoniae (K.p) and
S. marcescens (S.m) (1st site)
attTn7 Sites.

LIST OF TABLES

Chapter 1

- 1.1 Sequence Specificity for Insertion.....47
of Transposable Elements.
- 1.2 Effect of Base Substitutions on Tn5.....51
Target Duplication Sequence and Hot
Spot Selection.
- 1.3 DNA Sequence of Three Independent....54
Insertions of IS4 in attIS4 Sites
in galT.
- 1.4 DNA Sequence of Four Independent.....57
Insertions of Tn1545 in attTn1545
Sites.

Chapter 2

- 2.1 Antibiotic Resistance Markers for.....118
Different Bacterial Genera.

Chapter 3

- 3.1 Results of the M13 Transduction.....181
Frequency Assays.

Chapter 4

- 4.1 Results of the M13 Transduction.....199
Frequency Assays of Deletion
Mutants of 958 bp attTn7 Site.
- 4.2 Results of the M13 Transduction.....209
Frequency Assays on 70 bp Mini
attTn7 Site (MAS).

Chapter 6

- 6.1 One Letter Amino Acid Code.....264

6.2	A Comparison of Tn7 Target.....	269
	Duplication in <u>att</u> Tn7 Sites and Non- <u>att</u> Tn7 Sites.	

Chapter 7

7.1	Results of the M13 Transduction.....	283
	Frequency Assays on Altered Spacing (Δ 1LOOP, Δ 2LOOP, Δ 3LOOP) and Deletion Clones (MASCPL77 and Derivatives).	
7.2	Effect of (i) Altered Spacing, (ii).....	286
	Potential Cruciform Structure on the Transposition Frequency of Tn7 as Established by Using the M13 Frequency Transduction Assay.	

Chapter 8

8.1	Distribution of Mutants in the 70 bp....	319
	<u>att</u> Tn7 Site.	
8.2	Results of the M13 Transduction.....	323
	Frequency Assays of the Watson Strand (C to T Mutants) <u>att</u> Tn7 Site.	
8.3	Results of the M13 Transduction.....	324
	Frequency Assays of the Crick Strand (G to A Mutants) <u>att</u> Tn7 Site.	
8.4	Mutants in the 36 bp <u>glm</u> S Box with....	325
	Negative Phenotype.	
8.5	Mutants in the 36 bp <u>glm</u> S Box with....	326
	No Effect on Transposition Frequency.	

Chapter 9

9.1	Comparison of Different Amido-....	339
	transferases at the Active Site Amino Acid Residues.	
9.2	Restriction Pattern of Some of.....	350
	the Type I,II and III Enzymes.	

ABBREVIATIONS

$r(s)$	
Ap	Ampicillin resistant (sensitive)
AMV	Avian Myeloblastosis Virus
ATP	Adenosine triphosphate
<u>att</u>	Attachment site
<u>att</u> Tn	Attachment site of transposon
bp	Base pair
BSA	Bovine serum albumin
CCC DNA	Covalently closed circular DNA
cfu	Colony forming units
$r(s)$	
Cm	Chloramphenicol resistant (sensitive)
c.p.m.	Counts per minute
C-terminus	Carboxy terminal amino acids
d	Dalton
DNA	Deoxyribonucleic acid
DNAase	Deoxyribonuclease
dATP	Deoxy adenosine triphosphate
dCTP	Deoxy cytidine triphosphate
dGTP	Deoxy guanosine triphosphate
dTTP	Deoxy thymidine triphosphate
dNTP	Deoxynucleoside triphosphate
ddATP	Dideoxy adenosine triphosphate
ddCTP	Dideoxy cytidine triphosphate
ddGTP	Dideoxy guanosine triphosphate
ddTTP	Dideoxy thymidine triphosphate

DTT	Dithiothreitol
E	Exponential (e.g. 1×10^{-4} is written as $1E-4$)
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
F'	Fertility factor
$\frac{glmS}{r(s)}$	Glucosamine synthetase
Gm	Gentamycin resistant (sensitive)
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid
IPTG	Isopropyl-beta-D-thiogalactopyranoside
IR	Inverted repeat
IS	Insertion sequence
Kb	Kilobase pairs
Kdal	Kilodalton
MH	Mueller-Hinton
m.o.i.	Multiplicity of infection
mRNA	Messenger RNA
m.w.	Molecular weight
N	Any nucleotide
NaOAc	Sodium acetate
NH OAc	Ammonium acetate
N^4 -terminus	Amino terminal amino acids
O.D.	Optical density
ori	Origin of replication
ORF	Open reading frame
P	Promoter
PEG	Polyethylene glycol

PFU	Plaque forming units
PMSF	Phenyl methyl sulphonyl fluoride
POPOP	1,4-Bis(5-phenyl-2-oxazolyl)-benzene; 2,2'-p-phenylene-bis(5-phenyloxazole)
PPO	2,5-Diphenyloxazole
RE	Restriction endonuclease
RF	Replicative form (double-stranded) DNA of single-stranded bacteriophage
RNA	Ribonucleic acid
RNAase	Ribonuclease
RS	Recognition sequence
SDS	Sodium dodecyl sulphate
SSC (20X)	Sodium chloride (3M), Sodium citrate (0.3M)
TBE buffer (1X)	10.8g/l Tris base, 5.5g/l boric acid and 0.93g/l EDTA in distilled water
TCA	Trichloroacetic acid
<u>r(s)</u>	
Tc	Tetracycline resistant (sensitive)
TE	10mM Tris.HCl, pH 8.0; 1mM EDTA pH 8.0
TEMED	N,N,N',N'-Tetramethyl ethylene diamine
Tn	Transposon (N.B. underlines have been omitted from all transposons, e.g. Tn <u>7</u> is written Tn7)
<u>tns</u>	Transposition function
<u>r(s)</u>	
Tp	Trimethoprim resistant (sensitive)
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA

<u>Uas</u>	Upstream activation sequence
uv.	Ultra-violet radiation
X-gal	5-Bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Δ	deletion
ΔG	Thermostability value of secondary structure in RNA/DNA
::	Novel joints

CHAPTER 1

Introduction

1.1 Definition, History and Nomenclature of Transposable Elements:

Transposable elements are discrete segments of DNA which have the ability to alter their positions within the genome. Transposable elements encode proteins (e.g. transposases) required for their transposition, a recombination event which generally is independent of DNA sequence homology.

The first evidence of the presence of transposons was demonstrated by the excellent genetic experiments of Barbara McClintock on the common corn plant Zea mays who named them controlling elements (McClintock, 1951 and 1955). Since then, an extensive biochemical and genetical study has been made on a variety of these extraordinary elements.

Their existence has been documented in bacteria (Cohen, 1976), yeast Saccharomyces cerevisiae (Fink et al., 1980), fruit fly Drosophila melanogaster (Spradling and Rubin, 1981), nematode Caenorhabditis elegans (Emmons et al., 1983), vertebrates (Georgiev et al., 1984), slime mould Dictyostelium discoideum (Rosen et al., 1983) and of course

in plants (Doring and Starlinger, 1986; McClintock, 1951).

Clues as to the molecular mechanisms underlying DNA transposition only became clear from studies on genetic rearrangement caused by such elements in bacteria (Bukhari et al., 1977).

The primary goal of this work was to understand the transpositional properties of a bacterial transposable element, Tn7. Therefore, this introduction will be mainly confined to a discussion of the prokaryotic transposable elements. A number of reviews of prokaryotic transposable elements have been published in recent years which describe their properties in detail (Berg and Berg, 1983; Calos and Miller, 1980; Cohen and Shapiro, 1980; Grindley and Reed, 1985; Heffron, 1983; Kleckner, 1981; Kleckner, 1983; Toussaint and Resibois, 1983).

The gigantic (39 Kb) temperate phage, Mu, was the first transposable element to be described in prokaryotes (Taylor, 1963). It was shown later that during the lytic cycle, Mu transposes to many new locations and generates chromosomal rearrangements including deletions, inversions and duplications of E. coli chromosomal DNA (Bukhari, 1976).

A large number of polar mutations in the E. coli chromosome were identified in the late 1960's which were caused by the insertion of DNA sequences of variable length (Jordan et

al., 1968). The occurrence of different classes of these insertion sequences (IS) was later discovered throughout the E. coli chromosome and in many bacteriophages, plasmids and episomes (Malamy et al., 1972; Starlinger and Saedler, 1972). In the early 1970's it was reported that in many bacteria, antibiotic resistance genes were spread at high frequencies via transposable elements (Cohen, 1976; Hedges and Jacob, 1974).

Bacterial transposons have been shown to carry markers for heavy metal resistance (such as mercuric ion resistance of Tn21 and Tn501, Bennett et al., 1978), a heat stable toxin (Tn1681, So et al., 1979), genes for sugar utilisation (Tn931, Cornelius et al., 1978), genes for conjugative transfer (Tn196, Gawron-Burke and Clewell, 1982), and genes for sex pheromones (Tn918, Clewell et al., 1985).

Transposable elements which carry drug resistance, heavy metal resistance etc., in addition to their transposition genes are called transposons and those that carry only genes required for transposition are called insertion sequences (IS elements). Current interest in transposable elements centres on at least five different areas; (1) for the evolutionary theorist, their role in evolution by promoting DNA rearrangements within organisms and genetic flux between organisms, (2) for medicine, their role in the clinical spread of antibiotic resistance in many bacteria, (3) for the geneticist, the regulatory mechanisms that control their transposition, (4) for the molecular

biologist, a fundamental understanding of the actual mechanism(s) of DNA transposition. This involves an understanding of the enzymology of transposition at the level of the substrate requirements, the DNA-protein interactions and the changes in DNA topology that take place during transposition, and (5) finally, transposable elements have proved powerful tools for genetic analysis of the organism by transposon mutagenesis.

The basic features of transposable elements (Tn) include: inverted repeated ends, Tn encoded transposase(s) that recognises the ends and brings about transposition. Almost all transposable elements make a short duplication of the target sequence whose length is characteristic of a particular transposable element. A variety of mechanisms exist for transposition, i.e. replication with or without giving rise to a cointegrate (formed by Tn donor and recipient replicon fusion which has a transposon at each junction of target and donor DNA) and simple insertion by "cut and paste". The properties of various transposable elements are discussed in detail later in this chapter.

Transposable elements can be divided into three classes based on their structure and transpositional properties as described below. However, a large number of bacterial transposons such as Tn7, Tn917, Tn554 and Tn1545 do not fall into any of these classes (Grindley and Reed, 1985; Kleckner, 1981).

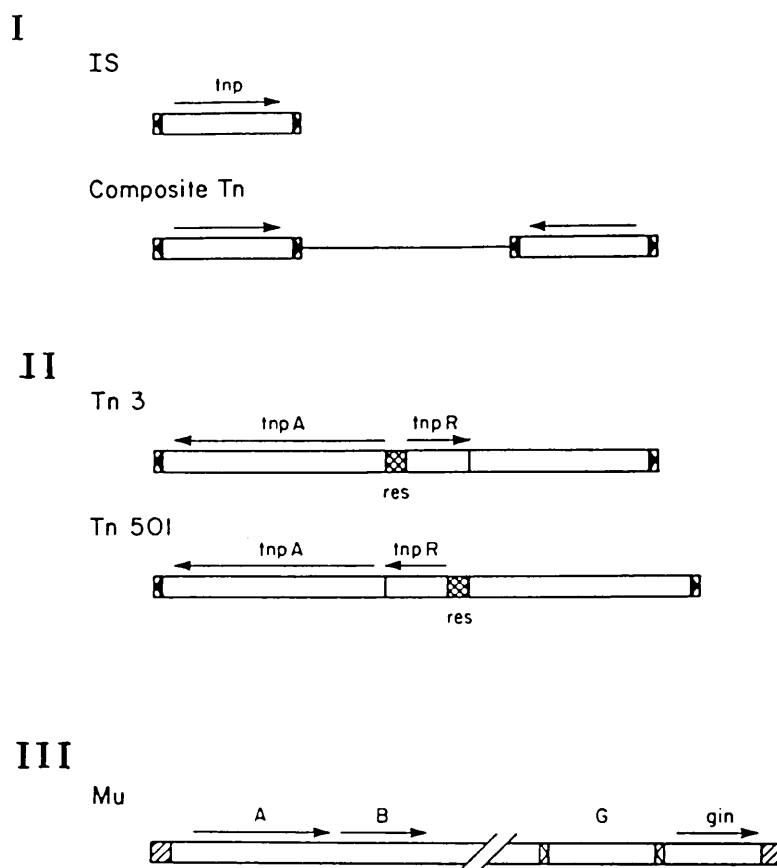


Figure 1.1 Three classes of transposable elements. Taken from Grindley and Reed (1985).

Class I includes insertion sequence elements (IS) and composite transposons. ISs range from 750 bp to 1600 bp in length and encode one, or sometimes two, proteins [transposase(s)] for transposition. Composite transposons consist of a DNA segment, encoding a phenotypic determinant such as antibiotic resistance, flanked by two copies of IS elements with direct or inverted orientation.

Class II transposons, also called the Tn3 family of transposons, have tnpA and tnpR genes which code for transposase and cointegrate resolvase respectively and also

contain a cis-acting res site required for cointegrate resolution. A cointegrate structure is a required intermediate in the transposition of the family of elements. Two sub-groups exist in this class in which the transcription of tnpA and tnpR genes are either in opposite orientation (Tn3) or in the same orientation (Tn501) and also their transposition functions (particularly tnpR gene function) are not completely related (Schmitt et al., 1979).

Class III includes bacteriophage Mu and related phages like D108 which have characteristics of transposons. These phages encode genes for replication and transposition. In phage Mu the gene products of the A and B genes encoded in the left (immunity) end are necessary for efficient transposition. In addition an interesting function gin is encoded from the right end of Mu. The gene product of gin is a site-specific recombinase which is required for site-specific inversion (flip-flop) of a 3 Kb DNA segment (the G segment) which determines the phage adsorption on two alternative bacterial hosts (Bukhari and Ambrosio, 1978). Interestingly, although the gin gene is not required for Mu transposition it is related to the tnpR gene product of Tn3 and Tn501 (Grindley and Reed, 1985).

1.2 Consequences of Transposition:

In prokaryotes DNA rearrangement in vivo has been shown to occur by reassorting the information in at least two ways; (i) by homology dependent recombination, and (ii) by non-homology dependent recombination which generally does not require extensive DNA homology and is also referred to as illegitimate recombination. The latter class involves site-specific recombination and transpositional recombination and generally takes place independent of recA function.

The process of transposition differs from site-specific recombination in the fact that in transposition, at some stage, DNA synthesis (either at the target site or the whole transposable element) is achieved (Grindley and Reed, 1985). The two apparent consequences of transposition associated with different transposable elements are discussed in sections 1.2.1 and 1.2.2.

1.2.1 DNA Rearrangements Associated with Transposable Elements:

All transposable elements promote intra- and inter-molecular DNA rearrangements in addition to simple insertion (Kleckner, 1981). Simple insertions can be achieved either by a conservative transposition in which the transposon is cut out from the donor replicon and pasted into the target site, or by a replicative transposition in which the transposon is duplicated and

then translocated into the target site (Derbyshire and Grindley, 1986; Grindley and Reed, 1985; Kleckner, 1981; Shapiro, 1983). Intra-/inter-molecular DNA rearrangements include; deletion inversion, replicon fusion (cointegration), deletion, excision, adjacent insertion and inverse transposition are generally associated with the transposable elements (Derbyshire and Grindley, 1986) (see figure 1.2). Inverse transposition is related to composite transposons in which the two IS elements flanked by a phenotypic determinant can promote the inter-molecular transposition of the donor replicon, instead of normal selectable determinant, into the recipient replicon (Kleckner, 1981). The intra-molecular transposition of composite transposons can also result in either adjacent deletion or deletion inversion, also called duplicative insertion (see figure 1.2). Any model of transposition mechanism must also be able to explain these DNA rearrangements.

Transposable element (IS element) mediated DNA rearrangements also include the integration of conjugative F and R plasmids into the bacterial chromosome upon transposition (replicon fusion) (Ohtsubo and Ohtsubo, 1977).

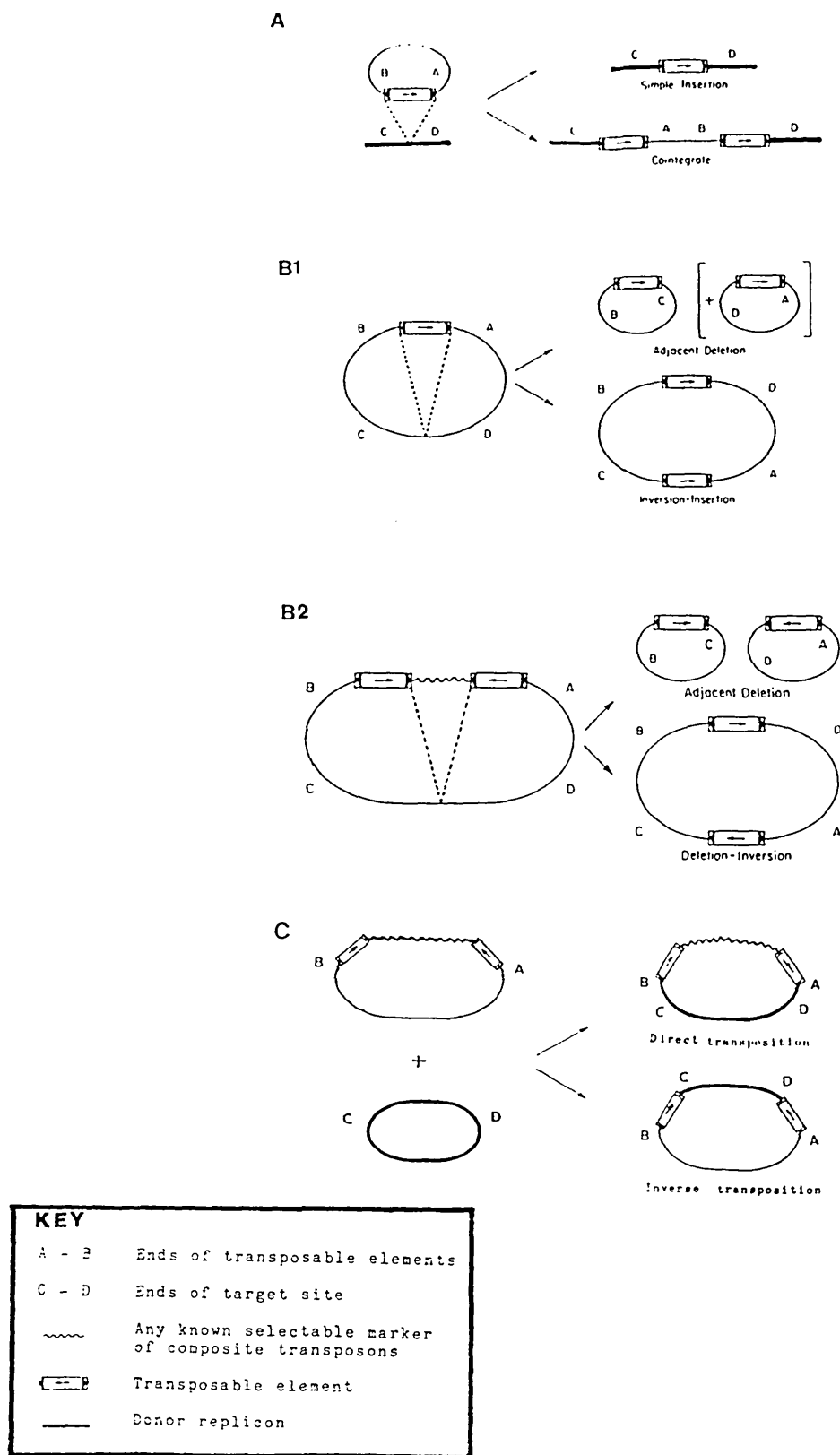


Figure 1.2 Types of transposable elements promoted events.
A Inter-molecular transposition of simple transposon.
B Intra-molecular transposition of simple and composite transposon shown as B1 and B2 respectively.
C Inverse transposition of composite transposon.

Since several copies of IS elements exist in F and R plasmid as well as in the chromosome, integration can also occur by homologous recombination involving two copies of an IS element, one from each replicon. This demonstrates the importance of the IS elements as substrates for homology dependent recombination, and any two copies of an element within the cell can participate in recombination. Depending on the relative orientation of the two copies of IS elements, intra-molecular reciprocal homologous recombination can result in either deletion or inversion of the adjacent sequence.

All transposition events involve the joining of transposon ends to the new sites in intra- or inter-molecular fashion. The inter-molecular events are of two types. The first type involves a direct transposition with or without duplication of the transposable element, (the conservative mechanism of transposition) of, for example, Tn5, Tn10, IS10 and IS50. The second type involves an indirect transposition achieved by replicon fusion with duplication of the transposable element, (the replicative mechanism of transposition) of Tn3, Mu, IS8, and IS102 (Grindley and Reed, 1985). However, some elements, such as IS1 and Tn903, appear to transpose via both mechanisms (Derbyshire and Grindley, 1986). The models for transposition will be discussed in some detail after the general characteristics of mobile elements on which they are based.

The consequences of inter-/intra-molecular transposition by a donor suicide "cut and paste" process include the excision of the transposable element from the donor site and insertion into the target site. This can result in a linearised donor replicon which may be degraded by exonucleases. Intra-molecular transposition by this mechanism can result in a circular product only if the target lies within the transposon (see figure 1.3). This is achieved when the ends of the elements are joined to the target in one of two possible orientations, which also generate deletions or deletion-inversions.

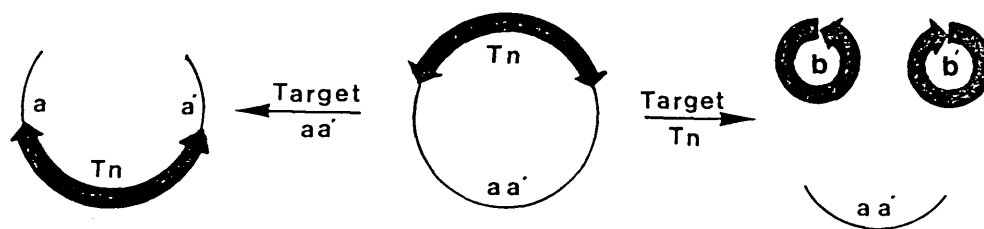


Figure 1.3 Intra-molecular transposition by a "cut and paste" mechanism. Transposition of a transposon (bb') by this mechanism results in a circular product (b) or (b') when the target lies within bb'. This leaves the donor replicon (aa') as a linear molecule (Derbyshire and Grindley, 1986).

1.2.2 Alteration of Gene Expression by Transposable Elements:

The insertion of transposable elements usually results in polarity; altering the expression of genes distal to the site of the insertion. Some IS elements generate efficient promoters upon insertion to certain new sites and some carry mobile promoters. For example, IS2 has been shown to enhance the expression of a class of ampC beta-lactamase gene in E. coli by 20% by creating a new, more efficient promoter upon insertion within the old promoter (Jaurin and Normark, 1983). Insertion of transposable elements can also lead to activation of downstream genes or operons. For example, insertion of IS2 and IS3 in the arg operon (Galansdorff et al., 1980) and IS1 and IS5 in the cryptic bgl operon (Reynolds et al., 1981) have been shown to activate these operons. A similar observation has also been made that insertion of IS2 in the int gene of bacteriophage lambda activates the expression of the int gene (Pilacinski et al., 1977).

The detailed mechanisms of gene expression in these systems has as yet not been elucidated. Transposable elements may interrupt the expression of genes by either directly inserting into the genes or by inducing polar effects on downstream genes in an operon. Transposable elements can produce transcriptional polar effects (premature termination of transcription) by introducing transcriptional terminators if contained within the

element. They can also produce translational polar effects by interrupting translation, thus causing RNA polymerase to terminate at cryptic rho dependent terminators within the coding region. In both cases premature termination of transcripts cause the loss of expression of distal genes.

1.3 Structural Organisation and Functions of Prokaryotic Transposable Elements:

Most transposable elements show near-perfect direct or inverted-repeats at their ends to which it is thought transposase binds. However, there are some exceptions; the ends of bacteriophage Mu share short regions of homology but do not form convincing inverted repeats (Kahmann and Kamp, 1979). Tn554 (Murphy and Löfdahl, 1984) and Tn1545 (Caillaud and Courvalin, 1987) are significantly different known prokaryotic transposons without any direct or inverted repeats at their ends.

The length of inverted repeats vary considerably from one transposable element to another. IS elements have short (20-40 bp) inverted repeats. Deletion at either end, or both ends, of a transposable element abolishes transposition. This was first demonstrated by Heffron et al. (1977) who showed that Tn3 ends can not be complemented in trans.

In composite transposons, a segment of DNA that encodes

resistance to antibiotics etc., is flanked by a pair of class I elements either in a direct or in an inverted fashion. These elements constitute the ends of composite elements. For example, IS50 in the case of Tn5, and IS1 in the case of Tn9, flank genes for resistance to kanamycin and chloramphenicol respectively. The left and right IS elements of composite transposons may be identical (IS903L is the same as IS903R in case of Tn903) or different (IS50L differs from IS50R of Tn5 by 1 base pair).

1.3.1 Transpositional Immunity:

Many transposons display a phenomenon referred to as transpositional immunity (Lee et al., 1983; Wallace et al., 1980). A plasmid already containing a copy of a transposon is immune to further insertion. Dissection of this phenomenon in Tn3, Tn554 and phage Mu has shown that only one copy end needs to be present in a plasmid to inhibit further insertion (Lee et al., 1983; Murphy, 1983; Toussaint and Resibois, 1983). It was found that the 38 bp region of the left or right inverted repeat of Tn3 exerts immunity over itself but not over the related transposons gammadelta or Tn501 (Arthur et al., 1984). For Tn3 the degree of transpositional immunity seems to be also influenced by the level of tnpA gene product (Heritage and Bennett, 1984). The mechanism of transpositional immunity has remained obscure until recently. It was shown using in vitro transposition system that an end of Mu on a potential

target replicon de-stabilises the pre-transposition complexes (Aduzama and Mizuuchi, unpublished results).

1.3.2 Target Sequence Duplication at the Sites of Insertion of Transposable Elements:

Transposition in both prokaryotes and eukaryotes is characterised by the generation of short (2-13 bp), direct duplications of the target DNA (Grindley and Reed, 1985; Herman and Shaw, 1987). Exceptions are IS200 of Escherichia coli (Lam and Roth, 1986), ISH1.8 of Halobacterium halobium (Schnabel et al., 1984), Tn554 of Staphylococcus aureus (Murphy and Löfdahl, 1984) and Tn1545 of Streptococcus pneumoniae (Caillaud and Courvalin, 1987) which do not generate any duplication of sequences upon insertion.

In most cases, the transposase is thought to create a staggered cut at the point of insertion and the ends of the transposon are ligated to the longer overhanging strands. Target duplication would then result when the gaps formed are repaired by DNA synthesis.

Most elements can be placed into two groups according to the number of base pair duplications. One causing 5 bp duplications and another causing 9 bp. The 5 bp group includes Tn3, Mu, gammadelta, Tn7 and the Ty element of yeast. The 9 bp group includes IS1, Tn5, Tn9 and Tn10

(Bukhari, 1981).

A large number of variant IS elements have been isolated which generate variable length duplications of the target DNA. The prototype IS1 and the IS1 flanked by transposon Tn9 generates 9 bp repeats during either transposition or cointegration. An IS1 element (called IS1K) with a single base pair change (G to T transversion) in the terminal inverted repeat sequence has been found which generates 8 bp target duplication (Iida et al., 1981). An IS1 element originating from bacteriophage P1 generates both 8 and 9 bp target duplication (Iida et al., 1985). Another IS1 variant IS1F was found in the chromosome of Shigella flexneri and causes 8 bp duplication (Ohtsubo et al., 1984). Iida and Hiestand-Nauer (1986) have reported a 10 bp duplication generated by an IS1 element but the sequence of this element is not yet known.

The DNA sequences of IS1K and IS1F reveal 7 and 74 nucleotide substitutions respectively in comparison to IS1R. It seems apparent that some of these mutations at the ends of the element or in a target determinant domain of the transposase are responsible for this variation. In order to test this hypothesis that base substitution in the IS1 sequence affects the length of the nucleotide sequence which is duplicated during IS1 mediated cointegration. Machida and Machida (1987) constructed a set of hybrid IS1s, between IS1R and IS1K, which have four base substitutions in the insA gene and three in insB gene of

the IS1 element. They found that IS elements with any of these mutations had caused a DNA duplication of variable length (7, 8 or 9 bp) at the target site. This observation has suggested that both genes are involved in the target duplications and all other IS variants may have naturally occurring mutations in these genes.

Two other IS elements IS3 and IS4 are also shown to generate variable duplications, 3-4 and 11-12 base pairs respectively (Habermann et al., 1979).

1.3.3 Transposable Element Encoded Transposition Proteins:

The comparative study of different elements suggests the need of at least one diffusible gene product (transposase) encoded by the transposable element for transposition. For example, the transposase of Tn3 is tnpA product and is approximately 113 Kdal. This protein has been over-produced and purified to homogeneity (Morita et al., 1987). In the Tn3 family, a second protein, resolvase, is needed to resolve the cointegrate molecule which is an essential intermediate in the pathway for transposition of these elements. The detailed work on proteins encoded by class II transposable elements indicates that transposase and resolvase function can be efficiently complemented in trans within the sub-classes of Tn3 family transposons.

The transposase recognises the termini of the element (the inverted repeat sequence) and mediates the formation of cointegrates of donor and recipient replicons separated by directly orientated copies of the element. The resolvase then catalyses the site-specific recombination at the res site of the element to give a new recombinant and the original donor, now both carrying the transposon (Heffron et al., 1979; Sherratt et al., 1981). The resolvase (tnpR gene product) is the repressor of tnpA gene and is under autogenous negative regulation. The resolvase protein from Tn21 (Halford et al., 1985, Tn3 (Krasnow and Cozzarelli, 1983) and gammadelta (Reed, 1981) have been cloned in expression vectors and purified to homogeneity.

A sub-set of class II elements, that includes Tn21 and Tn1721, differs in their resolvase functions from Tn3. For example, a cointegrate formed with a tnpR mutant of Tn21 is resolved in vivo by complementation with Tn501 or Tn1721 but not with Tn3. The predicted amino acid sequence (from the DNA sequence) of the resolvases of Tn21 and Tn501 have 153 out of 186 identical amino acids while only 56 out of 186 are identical between the resolvases of Tn21 and Tn3. This suggests an evolutionary divergence of the structure and action of the tnpR gene product.

The transcription of the tnpA and tnpR in Tn3, Tn1000 and Tn2501 is in opposite directions, while in Tn21, Tn501 and Tn1721 it is in the same direction (Schmitt et al., 1985). The position of the res site is also different in these

elements. When transcription of the tnpA and tnpR genes is in the same orientation, the res site is located adjacent to the promoter of the tnpR gene. If the transcription is in the opposite direction, then the res site is located in between the genes but again adjacent to the promoter of the tnpR gene.

The composite transposon Tn5 (5700 bp and kanamycin resistant) contains 1534 bp inverted repeats (IRs); the right IR encodes a functional transposase and the left contains an ochre mutant allele of the transposase gene (Auerswald, et al., 1980). Another well studied composite transposon Tn10 (9300 bp and tetracycline resistant) has 1329 bp inverted repeats (IRs) of an insertion element IS10. The right IR is fully functional, can transpose independently and encodes a single transposase protein which can act at the termini of IS10 to promote IS10 or at the termini of Tn10 to promote Tn10 transposition (Kleckner, 1983).

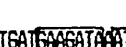
A Staphylococcus aureus transposon Tn554 (6600 bp and spectinomycin resistant) encodes at least three gene products [tnpA (43 Kdal), tnpB (74 Kdal) and tnpC (14 Kdal)] whose functions are required for efficient transposition as evidenced by deletion analysis of the transposon (Murphy et al., 1985).

In the case of Mu, two phage encoded proteins A (about 70 Kdal) and B (about 33 Kdal) have been shown to be required

for an efficient in vivo and in vitro transposition of phage Mu. Mu A is regarded as a transposase because it is essential for transposition and it specifically recognises the DNA sequence at the ends of the Mu genome (Craigie et al., 1984). Mu A gene product has been over-produced and purified to homogeneity (Craigie and Mizuuchi, 1985). Mu B has been shown to enhance the transposition in vivo and in a defined system in vitro. Recent evidence from in vitro work has suggested that Mu B is an ATPase and promotes inter-molecular DNA strand transfer (Maxwell et al., 1987).

1.4 Specificity for Insertion of Transposable Elements:
Among the varieties of prokaryotic transposable elements studied, the behaviour of insertion is surprisingly different from element to element. However, fine structure analysis and genetic mapping of a large number of insertions of different transposable elements suggest that most transposons insert with preference at some target sites. On the basis of insertional properties, transposable elements can be divided into three classes. The first class inserts randomly (phage Mu, IS1, Tn3), the second has a low degree of target specificity (Tn10, IS50) and the third with highly specific target preferences (IS4, ISH1, Tn7, Tn554, Tn1545).

Table 1.1 Sequence Specificity for Insertion of Transposable Elements.

Transposable elements	Consensus sequence at the insertion sites of transposons	Number of independent events studied	References
Group I (random transposable elements)			
IS1	A-T rich, preferential cleavage at G-C	7	Saedler <i>et al.</i> , 1981; Iida, 1982
Tn3	A-T rich	65	Tu and Cohen, 1980
Tn9	G-C pair preference at the ends of 9 bp target duplication sequence	23/28	Galas <i>et al.</i> , 1980
Tn5	G-C pair preference at the ends of 9 bp target duplication sequence	19/36	Berg <i>et al.</i> , 1984
Group II (region specific transposable elements)			
IS5	5'- ∇ C(T)A(G) ∇ -3' A A	3	Engler and Van Bree, 1981
Tn10	5'- ∇  -3'	11	Halling and Kleckner, 1982
Group III (highly specific transposable elements)			
IS4	5'-  CAN  -3' (2-3bp) (11-18bp)	3	Klaer <i>et al.</i> , 1980
ISH1	5'-  AGTTATTG  -3'	2	Simsek <i>et al.</i> , 1982
Tn554 (primary attTn554 site)	5'-GAT   ATAACTATG-3'	3	Murphy and Lofdahl, 1983
Tn554 (secondary attTn554 site)	5'-TTT   -3'	2	Murphy and Lofdahl, 1983
Tn1545	5'-   -3' (3 to 5bp)	4	Cailland and Courvalin, 1987
Tn7	5'-GCGGG-3' in the transcriptional terminator of glsS of attTn7 site (958 bp EcoRI fragment) from <i>E. coli</i> chromosome.	2	Lichtenstein and Brenner, 1982; Gay <i>et al.</i> , 1986

∇ Denotes the point of insertion of transposable elements.

Boxes show the region of homology between the primary and secondary attTn554.

Boxes with half arrows show the symmetrically disposed sequence at the insertion sites of transposable elements.

In the following section of this chapter, the insertional properties of some elements belonging to these three classes are discussed in some detail. Table 1.1 summarises the sequence specificity for insertion of some transposable elements.

1.4.1 Random Insertion of Transposable Elements:

An extensive study of targets for phage Mu, IS1, Tn3, Tn9 and Tn5 transposition have been shown to contain no apparent consensus sequence at their target sites.

The phage Mu was first to be shown to insert promiscuously into the lacZ gene at 75 different positions and all these sites were separable from each other by recombination (Bukhari and Zipser, 1972). The nucleotide sequence analysis of four independent insertions of Mu reveals no convincing homology at the insertion point (Kahmann and Kamp, 1979).

Tu and Cohen (1980) have sequenced 65 insertions of Tn3 into a small plasmid. A 4 fold preference for an A-T rich segment was detected within which multiple insertions at the same nucleotide position were found. One hot spot of this type had 26 occurrences and the respective 5 bp repeat was mostly A-T base pairs.

Preference for insertion of IS1 elements also seems to depend on the same criteria. However the consensus of the

9 bp duplication of the target site contains a G-C base pair either at both ends or at least at one end of the insertion. Very near to the site of insertion a similar observation was made for IS2. It inserts at nine different sites within a 1.7 Kb segment of phage P1 DNA which was moderately A-T rich (Arber et al., 1980). In the cases of IS1 and Tn3 their ends were found to have some homologies to their respective target repeats which suggests that such homologies may be involved in site selection during insertion.

What determines this preference for the insertion of elements discussed above is not clear. It may be the exposure of the particular region on the recipient plasmid in contrast to other inaccessible regions to the transposase of the elements. The insertion of these elements into A-T rich regions suggests a requirement for easy denaturability of DNA at the target site. This is also characteristic of RNA polymerase binding sites and apparently reflects the need for denaturing DNA during transcription initiation.

In another report Galas et al. (1980) have observed that 28 independent insertions of Tn9 occur in the final 160 bp fragment of the lacZ gene. The DNA sequence of 23 different insertion points have revealed a strong preference for a G-C base pair at each end of the 9 bp sequence duplicated during transposition. Twenty three out of twenty eight insertion events which resulted in a G-C

base pair at each end of the 9 bp repeat, can occur randomly with a probability of less than $1E-9$. Therefore this insertional preference has to be considered significant.

A similar observation was made in the case of Tn5 (Berg et al., 1983; Bossi and Ciampi, 1981). Bossi and Ciampi (1981) analysed the DNA sequence of three independent insertions of Tn5 in the histidine operon of Salmonella typhimurium. They found that all three insertions occur adjacent to G-C rich regions without any apparent sequence homology with each other. Berg et al. (1983b) performed the genetic mapping studies of 150 insertion sites of Tn5 in plasmid pBR322 and found that 75 insertions have occurred at 28 sites in the tetracycline resistance gene of the plasmid. Out of these insertions, 21 occurred at one site (site I) while 4-10 occurred at four other sites (sites II, III, IV and V). They called these 5 sites of insertion as hot spots. The DNA sequence comparison of these hot spots revealed that a G-C base pair is always present at the ends of the segment duplicated by insertion and other nucleotides seem to have little in common. In addition, of 28 low frequency insertion sites observed, half contained G-C at both ends and half contained G-C at only one end; none of these contained AT at both ends. Thus the presence of at least one G-C base pair seems to be of general importance for insertion of Tn5.

To test the importance of terminal G-C base pairs in the 9 bp duplication sequence in hot spot I, Berg et al. (1984) constructed pBR322 derivatives with a transition of G-C to A-T in the 9 bp target site and tested the effect of these base substitutions on the hot spot selection for Tn5 transposition (see table 1.2).

Table 1.2 Effect of base substitutions on Tn5 target duplication sequence and hot spot selection. Tn5 insertion represents an independent transposition of Tn5 from the chromosomal lacZ gene to the tet gene of pBR322 (wild-type or mutant) plasmid.

Mutant number	Target sequence Wild-type and mutants	Fraction in hot spot/tet ^s		
		I	II	Total
	1 9			
Wild-type	TAA GCTTTAATG CG	40	17	150
1	... A..... ..	0	7	62
2	T ..	13	8	66
3	C ..	15	6	97
4	G..... ..	5	0	33
5	A..... ..	3	5	27

Their studies concluded that mutation of G-C to A-T at position 1 of the target duplication sequence drastically reduces the hot spot usage. However the transversion from G to C in this position results in mild (about 2 fold) reduction in hot spot usage. While G to T or G to C mutation at position 9 did not severely affect the hot spot activity. This indicated that only one G-C bp may be important in position 1 in the recognition of the sequence. Since there are many similar positions in pBR322 that are

not used for insertion hot spots, it is possible that additional factors in the target DNA may guide Tn5 to its preferred site. Berg et al. (1984) also reported the insertion of nucleotides upstream to position 1, indicated a drop in hot spot usage, which suggests that some additional signal may be present upstream to position 1 in hot spot I. Whether, it is only the G-C base pair that governs the choice of specificity of Tn5 insertion in site I of plasmid pBR322 or others factors, it is not clear and need to be clarified.

This type of specificity is similar to that of several topoisomerase (Morrison and Cozzarelli, 1979). For example, DNA gyrase makes a four base pair staggered cut such that one strand is cut between a T-G dinucleotide.

1.4.2 Region Specific Insertion of Transposable Elements:

Two bacterial transposable elements (IS5 and Tn10) have been shown to have a defining property to insert repeatedly into a preferred consensus sequence (Engler and van Bree, 1981; Halling and Kleckner, 1982). A simple 4 bp duplication (5'-CTAG-3') target sequence consensus for three IS5 insertions was found (Engler and van Bree, 1981). Halling and Kleckner (1982) were able to identify a 6 bp symmetrical consensus sequence for 11 independent Tn10 insertions in the histidine operon of E. coli. This

consensus sequence (see table 1.1) was located within the 9 bp target duplication that is made by Tn10 during transposition. The ends of Tn10 and the consensus target DNA have some plausible homology as seen in the cases of IS1 and Tn3.

The juxtaposition of the recognition and the cleavage site in accordance with the symmetry of the perfect consensus sequence suggests that the target DNA may both be recognised and cleaved by sub-units of IS5 and Tn10 transposases. Hence this type of specificity is determined as for type II restriction endonucleases (Kessler and Holtkle, 1986). In type II restriction endonucleases the homologous palindromic recognition sequences are arranged from the shortest central palindromes AT, GC, CG and TA. By extending one nucleotide A to G to C to T on the left side and the complementary nucleotides on the right side would generate a tetranucleotide palindrome which is the minimum recognition sequence required by type II restriction endonucleases. The palindromic sequence requirement can be extended to 6 or 8 bp.

However, whether the presence of a symmetrical sequence alone presents an efficient target for IS5 and Tn10 integration has to be determined.

1.4.3 High Specificity Insertion of Transposable Elements:

Among the transposable elements studied so far IS4 (Habermann et al., 1979; Klaer et al., 1980), ISH1 (Simsek et al., 1982), Tn554 (Murphy and Löfdahl, 1984), Tn1545 (Caillaud and Courvalin, 1987) and Tn7 (Lichtenstein and Brenner, 1981; 1982) are reported to be highly specific in their insertion behaviour. The transpositional properties of these transposable elements are discussed below except those of Tn7 which are discussed in section 1.10.

IS4 which naturally exists in one or two copies in the E. coli chromosome seems to insert preferentially at a single site within the galT gene (Habermann et al., 1979). IS4 is presently known to insert in three different unique sites in several strains of E. coli K12. The DNA sequence comparison of these attIS4 sites shows that IS4 generates 11 or 12 bp duplication at the target site upon insertion (Klaer et al., 1980) (see table 1.3).

Table 1.3 DNA sequence of three independent insertions of IS4 in attIS4 sites in galT. The (■) denotes the point of insertion of IS4 which is variable (11 bp or 12 bp).

<u>attIS4</u> Site-1	5'-AAACAC■CAGAACCAAGAT■GCCTATTTTCGTTT-3'
<u>attIS4</u> Site-2	5'-AAACGC■CAGCCGCCGCCG■CGTTTTTTCTCGC-3'
<u>attIS4</u> Site-3	5'-AAACCTG■CACCCATGGGT■ACGTTTTCCCCAG-3'

The 11 bp target duplication sequence of Site-3 shares a CA nucleotide at one terminus of 12 bp of target duplication sequence observed at the other two insertion sites. In all three attIS4 sites a region of homology can be found which is symmetrically disposed 3 to 11 bp outside the region of target duplication sequence. The consensus sequence from this analysis expands to 31 bp (see table 1.1) and it is not known whether these sequence similarities have any relationship to the specificity of integration. Since the consensus sequence itself has not been tested as an efficient target for IS4 transposition, it is difficult to assess the role of these inverted repeats in the attIS4 sites.

ISH1, a transposable element found in Halobacterium, has been shown to insert specifically in the bacteriorhodopsin gene (BR) and the DNA sequence of two insertions reveal an 8 bp target sequence (5'-AGTTATTG-3') which does not contain a symmetrically disposed region. However, the target sequence from BR gene is flanked by an almost perfect 9 bp repeat (see table 1.1) and six nucleotides of these repeats align with nucleotides within the ends of ISH1 (Simsek et al., 1982).

Tn554 of S. aureus and Tn1545 of S. pneumoniae have many unique features that distinguish them from other prokaryotic and eukaryotic elements. These elements do not

contain the inverted termini, do not cause any base pair duplication of the target sequence and are highly site-specific for insertion.

Tn554 has been shown to transpose into the chromosome at a unique site and in one orientation. When the target site from the chromosome was cloned into a plasmid, the specificity of insertion was retained but orientation specificity was lost. Murphy and Löfdahl (1984) compared the insertion sites of Tn554 in plasmids with or without the target sites. Three independent insertions in the cloned target site and two insertions elsewhere in the plasmid were compared for DNA sequence homology. All hot-site attTn554 (primary) insertions were found within the sequence 5'-GAT GTA-3', however, Tn554 insertions in non-attTn554 sites (secondary) were found in a different sequence. Comparison of the primary and secondary sites revealed a region of homology 30 bp adjacent to the 3' of the insertion site of Tn554 (see table 1.1).

The ends of Tn554 are highly AT rich, 80% in comparison to 65% of S. aureus DNA. However, the AT contents of the primary and secondary att sites are variable, 46% and 76% respectively. This may suggest a possible role of AT content in the attTn554 site selection as the AT rich DNA is highly recombinogenic.

Tn1545 is a conjugative shuttle transposon and has been shown to transpose in both gram negative and gram positive

bacteria (Caillaud and Courvalin, 1987). It does not contain inverted repeats but contains three types of internal direct repeated sequences 5'-TAAAAATA-3', 5'-GCATAAAA-3' and 5'-AATTTACTACT-3' which are present asymmetrically. Four Tn1545 transpositions were reported in a consensus sequence 5'-TTTNT -(3 to 5 bp)-TAAAA 3' (see table 1.4.).

Table 1.4 DNA sequence of four independent insertions of Tn1545 in attTn1545 sites. Dots are introduced to align the consensus sequence marked by inverted arrows while (■) denotes the insertion point of Tn1545.

```

-----
attTn1545 Site-1      5'-CCTTTTTTTATT■ATTAAAAATCA-3'
attTn1545 Site-2      5'-TCTTTTATAG.■.ATAAAAATAG-3'
attTn1545 Site-3      5'-CACTTTCTCCA■..TAAAAAATA-3'
attTn1545 site-4      5'-ATGTTTATCAG■..TAAAAAAT-3'
-----

```

The consensus sequence obtained from these four insertion sites is highly AT rich and forms a secondary structure. The right part of the consensus has homology to the first internal repeat (5'-TAAAAATA-3') present at one end of Tn1545.

The very high AT rich regions of the attTn1545 sites and Tn1545 again suggests a contribution in high efficiency transposition as observed in the cases of IS1 and Tn3. A similar situation is also observed for eukaryotic topoisomerase 1, in which case a hexadecamer consensus

sequence motif (5'-AAACTTAGAAAAAAAAA-3') from the non-transcribed sequence of the ribosomal DNA was found to have a strong preference for topoisomerase 1. This consensus sequence is highly AT rich and a single base pair substitution in the recognition sequence reduces the extent of cleavage by 95% (Thomsen et al., 1987).

For target site selection of IS4, ISH1, Tn554 and Tn1545 the transposase of these elements seems to be an analogue to those restriction enzymes whose cleavage sequences are always located a certain number of nucleotides away from the DNA recognition site (Bickle, 1982; Boyer and Roulland-Dussoix, 1969; Kessler and Holtkle, 1986).

For example, Tn554 target recognition seems to be like MnII restriction enzyme in which case the enzyme recognises a sequence (5'-CCTG-3') and cleaves 7 nucleotides away from the 3' end of this sequence thus leaving a flush ended DNA at the cleavage site (Kessler and Holtkle, 1986). This analogy may be of superficial nature as the extent of DNA sequence required in any of the att sites is not known. Despite rather extensive studies on the nature of insertion sites of different transposable elements, it seems premature to make any conclusion on the requirements of the DNA sequence for insertion of any transposable element.

1.5 Control of Transposition:

Transposition processes are regulated by different mechanisms in various elements to achieve a balance between propagation of the element and viability of the host. The control of transposition of Tn3, Tn21, Tn5 and Tn10 has been elucidated in some detail and is discussed in this section.

In Tn3 the binding of resolvase (tnpR gene product) to the res site simultaneously suppresses the transcription of the transposase (tnpA gene product) and divergent transcription of the tnpR gene. Tn21, another Tn3 family transposon, encodes an additional 12.6 Kdal modulator protein which is thought to enhance transposition by acting on the tnpA promoter and suppressing the resolution by acting on a putative DNA site containing a large number of inverted repeats (Hyde and Tu, 1985).

A potential binding site for this modulator was located in the 5' region of tnpR with $\Delta G = -10$ Kcal per mol. Another putative site of action with $\Delta G = -40$ Kcal per mol was found only 95 bp upstream of the promoter of tnpA -35 region. Both these inverted repeated sequences are reminiscent of the trp attenuator sequence configuration. Tn501 transposition and cointegrate resolution was also shown to be regulated by the tnpM gene in in trans complementation experiments (Hyde and Tu, 1985).

Tn10 has 1329 bp inverted repeats of IS10 on either end. IS10R has been shown to contain the transcription start point for the transposase gene promoter, pIN at 81 bp and an opposing promoter pOUT at position 116 which provides the transcription of adjacent sequences after transposition insertion of Tn10 (Ciampi et al., 1982; Davis et al., 1985). Transposition of Tn10 is shown to be negatively regulated by two well defined systems (1) by an anti-sense RNA complementary to the transposase mRNA and (2) by DNA methylation. IS10, which is flanked by Tn10, controls the expression of its own transposase at the translational level.

This phenomenon, called multicopy inhibition, occurs by direct pairing between the transposase mRNA and small complementary regulatory RNA initiated at the IS10 pOUT promoter (see figure 1.4). The DNA sequence IS10 contains two dam methylation sites (5'-GATC-3'), one within the transposase promoter pIN and the other within the terminus where the transposase presumably binds. The methylation of these two sites blocks the transposition of Tn10 as observed with an increase in transposition of Tn10 in dam strains (Roberts et al., 1985). When IS10 is hemimethylated, pIN activity transposase is at least 1000 times higher than the fully methylated IS10. This suggests that pIN and the inner terminus do not function independently and that IS10 is activated to transpose following passage of a replication fork across IS10.

An increase in the transposition of Tn903 and Tn5 was also observed in *dam*⁻ strains. However, the transposition of Tn5 is also regulated by an inhibitory protein expressed from part of the transposase gene (Isberg et al., 1982; Johnson et al., 1982). Tn5 includes two insertion sequences, IS50L and IS50R, which flank a region of several antibiotic resistance genes (Auerswald et al., 1980; Berg, 1977).

IS50R produces at least two proteins, p1 and p2, which are encoded from the same reading frame. Truncated versions of these proteins, p3 and p4, are produced by IS50L which differ from IS50R at a single position near the 3' terminus of the reading frame (Rothstein and Reznikoff, 1981) but only p1 and p2 appear to have functional significance. The transposase protein p1 is required for transposition and is synthesised in a greater amount in the cell (Isberg et al., 1982). The inhibitor protein p2 lowers the transposition frequency when it is produced in a high amount.

The control of Tn5 transposition is more complex because overlapping promoters independently control expression of the transposase and inhibitor proteins. However, this does not involve the regulation of IS50R protein synthesis. The low level of expression of transposase from transcripts originating outside IS10 is believed to be predominately a result of secondary structure that masks the translational initiation region (Krebs and Reznikoff, 1986).

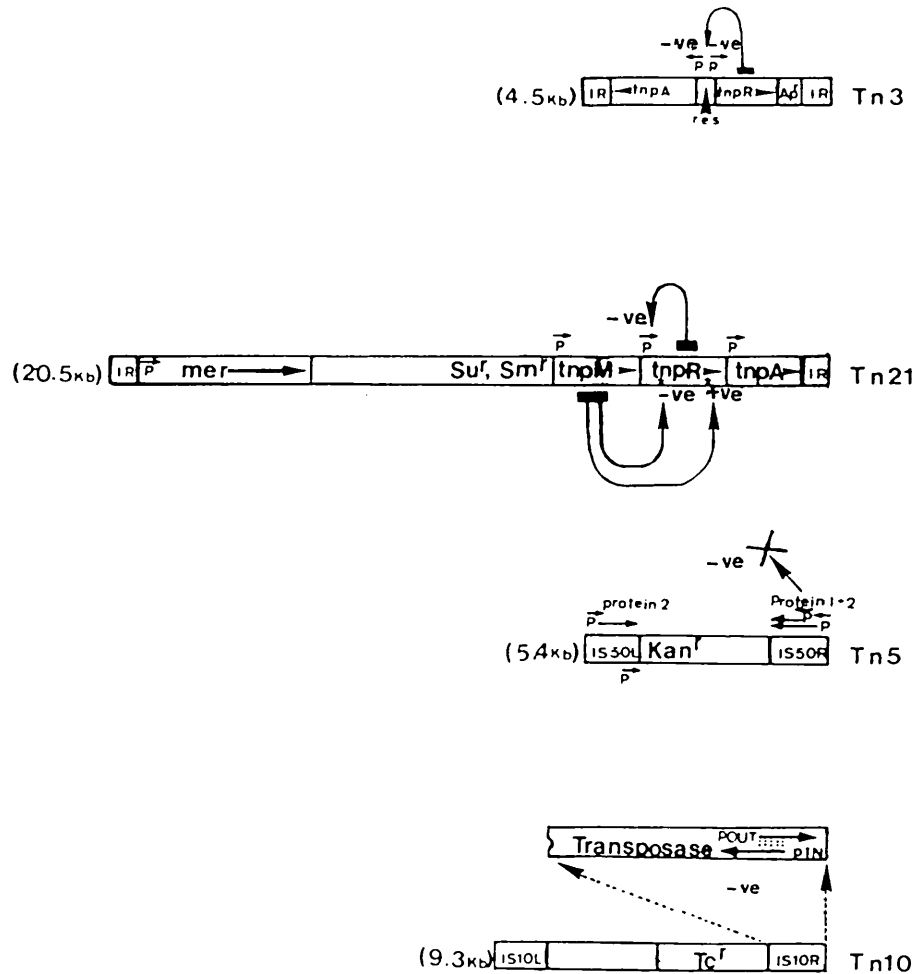


Figure 1.4 Regulatory mechanisms used by four transposons to control their transposition. In the figure IR represents the inverted repeats of the transposable element. In Tn3 the resolvase (product of *tnpR*) represses the transposase transcription as well as its own transcription. In Tn21, the modulator (product of *tnpM*) enhances the transposase function (+ve regulation) and suppresses the cointegrate resolution (-ve regulation). In Tn5, an inhibitory protein is translated in the same reading frame as the transposase and regulates transposition by directly binding to the transposase or transposase binding site. In Tn10, an anti-sense mRNA transcribed from pOUT inhibits the translation of transposase mRNA. In addition Tn10 transposition is also regulated by a complex mechanism of methylation as discussed in the text.

The presence of secondary structure that includes the translational initiation region of the transposase, as well as the transcriptional start site for the transposase promoter, appears to be the feature of other transposable elements including IS3 (Timmerman and Tu, 1985), IS5 (Kröger and Hobom, 1982) and IS30 (Dalarymple et al., 1984). Termination of read-through transcripts may also play a role in the control (Davis et al., 1985; Halling and Kleckner, 1982)

In Drosophila melanogaster the transposition of P element has been shown to be regulated in the germ line by precise excision of the third intron of the P element transposase (Laski et al., 1986). A developmentally regulated transcript of the TDD1 element in slime mould Dictyostelium discoideum is also reported (Rosen et al., 1983).

1.6 Models For Transposition:

A central dilemma exists concerning whether bacterial transposable elements replicate themselves in the process of transposition to generate two copies of the transposons, or if they excise from the old site and integrate into the new site without replication. Although there is much evidence supporting either a conservative mode or a replicative mode of transposition for certain elements, only in the case of Tn3, Tn10 and phage Mu has transposition been demonstrated by in vitro studies.

Most transposable elements are thought to involve either one or other of the mechanisms of transposition. Figure 1.5 shows the basic outline of two modes of transposition. Some elements such as IS1 and Tn903 appear to have a dual mode of transposition (Derbyshire and Grindley, 1986). In the following section, evidence is presented in favour of replicative and conservative modes of replication.

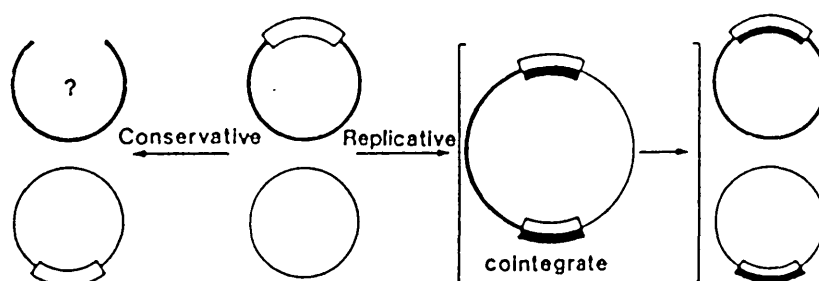


Figure 1.5 Replicative and conservative transposition. The transposon is shown as an open box and the donor and target replicons are represented by thick and thin lines respectively. Replicative transposition could proceed via a cointegrate formation or by simple insertion. Conservative transposition has several fates for the donor molecule (Derbyshire and Grindley, 1986).

1.6.1 Replicative Mode of Transposition:

Several models describing this mode of transposition have been proposed and they differ in several ways. In one group (replication first models), it is proposed that replication

of the element takes place before the breakage and rejoining of its ends to the target sequence (Kamp and Kahmann, 1981; Read et al., 1980). In another group (recombination first models) the introduction of staggered nicks at the target sequence occurs first, followed by the joining the ends of the element which initiates the duplication of the whole element (Arthur and Sherratt, 1979; Harshey and Bukhari, 1981). The recombination first model can be further sub-divided into symmetric and asymmetric models. The symmetric model predicts that both ends of the elements are ligated (Arthur and Sherratt, 1979) whereas in the asymmetric model only one end of the element is ligated to the target sequence at the initiation step (Galas and Chandler, 1981; Grindley and Sherratt, 1978). A cointegrate is necessary in symmetric models, whereas it is a by-product in asymmetric models (see figure 1.6).

A replication first model was originally proposed for phage Mu (Toussaint et al., 1977). The recombination first symmetric model was proposed for transposition of gammadelta and Tn3 (Arthur and Sherratt, 1979; Heffron et al., 1979; Reed, 1981).

Inter-molecular transposition of Tn3 and related transposons has been shown to proceed in two stages. Through the action of the tnpA gene product, donor and target replicons fuse to form a cointegrate in which the

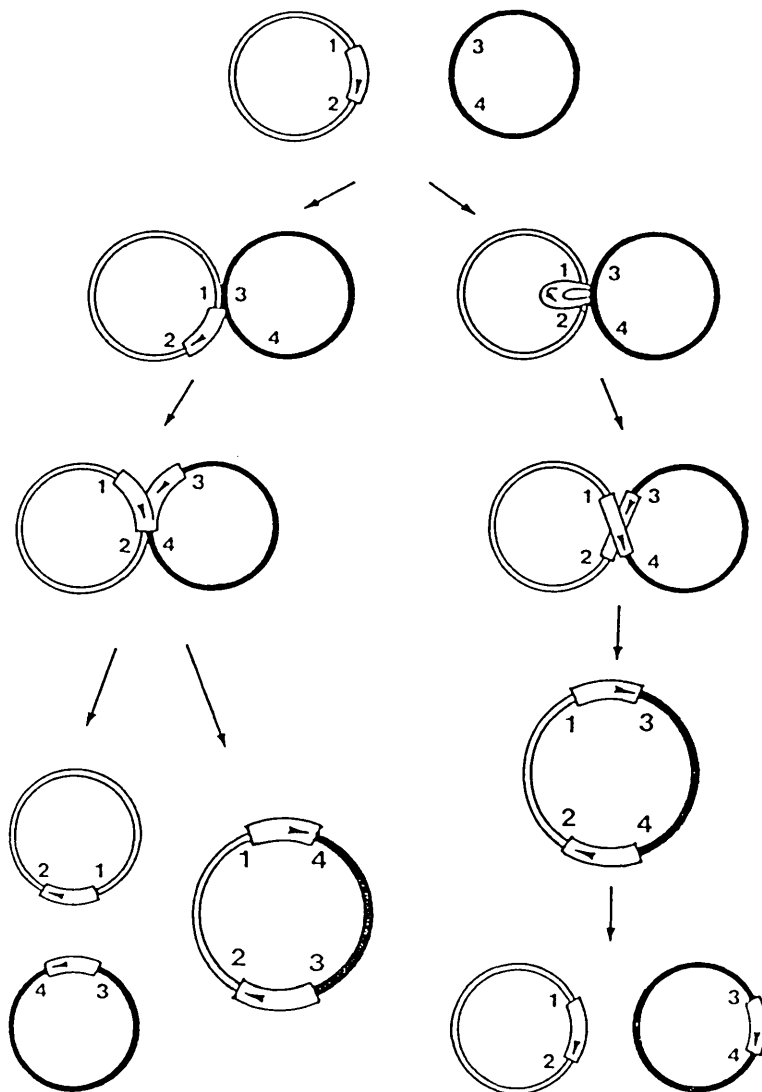


Figure 1.6 Simplified version of symmetric and asymmetric models of transposition. In symmetric model (right hand progression), ligation of both ends of transposable element with the target sequence initiates the duplication of the element to yield a cointegrate as an intermediate (Arthur and Sherratt, 1979; Shapiro, 1979). In asymmetric model (left hand progression), ligation of one end of transposable element with the target sequence initiates the duplication of the element (Galas and Chandler, 1981; Harshey and Bukhari, 1981). The transposable element is shown in a box and left and right inverted repeats at the donor replicon are marked by 1 and 2 respectively, while 3 and 4 denote the target site of the transposon. The arrows show the orientation of transposon.

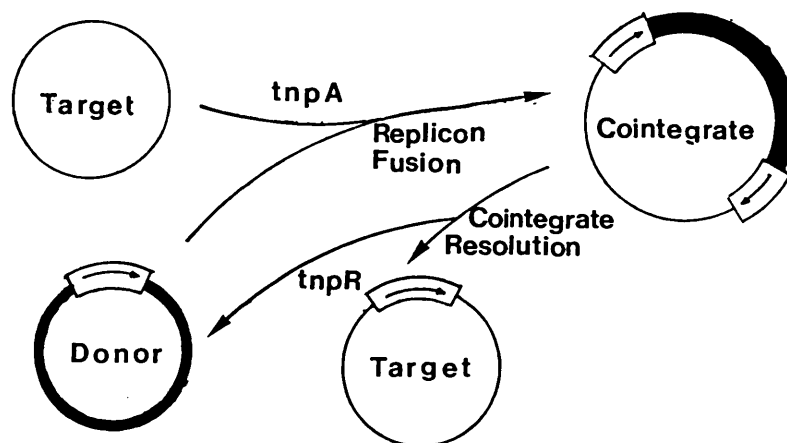


Figure 1.7 Cointegrate formation and resolution of Tn3. Donor and target replicons are fused in a replicative process by the action of transposase (tnpA gene product) to form a cointegrate which is resolved by the action of resolvase protein (tnpR gene product) into its constituent replicons: the target with inserted transposon and the regenerated donor. The transposon is shown in a box and the arrow shows the orientation of inserted transposon.

two copies of the transposon are orientated directly. This complex is resolved by the action of the tnpR gene product (site-specific recombination) at the res site in the transposon to produce the recipient and donor replicon with directly oriented copies of the transposable elements (Arthur and Sherratt, 1979).

It has been shown that approximately 30-45 bp of each terminal repeat of Tn3 are required for cointegrate formation. However, in the case of Tn21 and Tn1721 transposons deletion at one end can still form the cointegrates but the transposition is reduced to 5% (Avila

et al., 1984). Cointegrate resolution has been studied in vitro to elucidate the details of many different intriguing biochemical interactions between the resolvase and the res site (Krasnow and Cozzarelli, 1983; Reed and Grindley, 1981) some of which are discussed in section 1.5.3.

The early observation made by Ljungquist and Bukhari (1977) that phage Mu DNA replicates in the lytic infection has provided another evidence for the replicative mode of transposition. This was further elucidated by Mizuuchi (1983) who showed that cointegrate intermediates are formed in vitro through the action of Mu A, Mu B proteins and the E. coli histone-like protein Hu. Recent evidence suggests that Mu B protein is an ATPase and it facilitates the interaction between potential target sites and Mu A bound to the pair of Mu ends (Maxwell et al., 1987). They have also shown that with Mu A plus Mu ends plus target DNA, in the absence of Mu B and or ATP, the inter-molecular Mu DNA strand transfer reaction is extremely inefficient.

1.6.2 Conservative Mode of Transposition:

The first evidence of conservative transposition was shown by the simple insertion of phage Mu into the E. coli chromosome (Akryod and Symonds, 1983; Liebart et al., 1982). This observation was confirmed by Harshey (1984) who used dam methylation and the restriction enzymes DpnI (sensitive to dam) and MboI to differentiate between the

infected methylated Mu and its replicated forms, hemimethylated and unmethylated in a dam⁻ host strain. Upon integration, Mu retained the sensitivity to DpnI indicating it was fully methylated. The hemimethylated and unmethylated forms appeared only after a lag as expected. A model of direct non-replicative transposition was proposed by Berg (1977) and Berg et al., (1984) to explain Tn5 transposition (see figure 1.8).

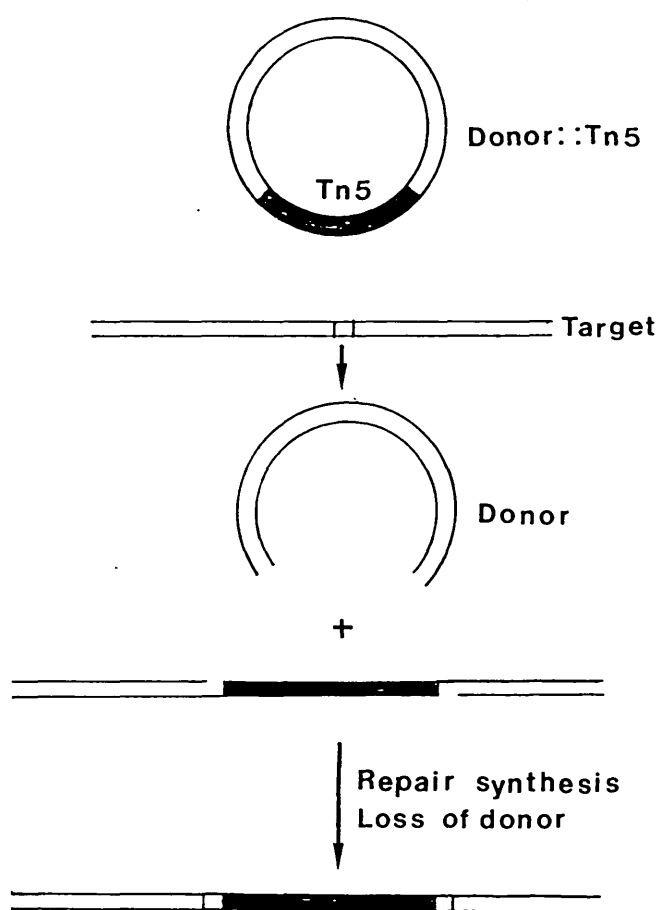


Figure 1.8 Model of conservative transposition of Tn5.

The transposition of Tn10 (IS10) has also been shown to occur by a conservative process (Bender and Kleckner, 1986). They constructed a Tn10 heteroduplex element with three single mismatches, one in the lacZ gene which was cloned in Tn10 and two close to the ends of Tn10. Tn10 was introduced into E. coli through a defective non-replicating lambda phage and tetracycline resistant colonies resulting from Tn10 transposition were selected on plates containing a beta-galactosidase indicator, X-gal. Sectorial blue and white colonies appeared which had resulted from a single transposition event (a simple insertion) as the mismatch, only 70 bp away from one end, was not repaired. This result indicated that Tn10 has inserted conservatively and that any repair replication at the target-transposon junction did not generally proceed as far as 70 bp from the transposon end.

Like Tn10 and initial insertion of Mu, the mechanism of transposition of Tn554 and Tn1545 is thought to be conservative but in the latter cases a double-stranded transfer of these elements to a flush, double-stranded break at the target site is implied.

In the cases of phage Mu, IS1 and IS903 where cointegrates are observed, there is also evidence that transposition could be conservative; the key question is, what governs the choice between replicative and conservative pathways? This may depend upon the state of the DNA (superhelicity,

methylation etc.) and the involvement of host or transposon encoded factors. For example, the infecting phage Mu DNA is effectively linear but is held in a non-covalent circle by an approximately 64 Kdal protein (Harshey and Bukhari, 1983) and the transposition is conservative. However, in the in vitro studies, Mu DNA is part of a superhelical molecule and the transposition intermediate formed indicates a replicative transposition.

1.7 Site-specific Recombination:

As mentioned in section 1.2 of this chapter, DNA rearrangements can be generated by recombination between homologous or non homologous DNA which involves a cross-over by a specific breakage-reunion reaction between two short regions of DNA. Site-specific recombination systems involve a site-specific recombinase and the recombining sites on which the enzyme acts preferentially. The best characterised recombination systems are integration and excision of phage lambda in the E. coli chromosome (Herskowitz and Hagen, 1980; Nash, 1981) and the resolution of cointegrates mediated by the transposon Tn3 and gammadelta (Arthur and Sherratt, 1979; Sherratt et al. 1980).

However there are many other site-specific recombination systems which vary in the requirements for the recombination reaction. They involve resolving of plasmids multimers (Summers and Sherratt, 1984) and transcriptional

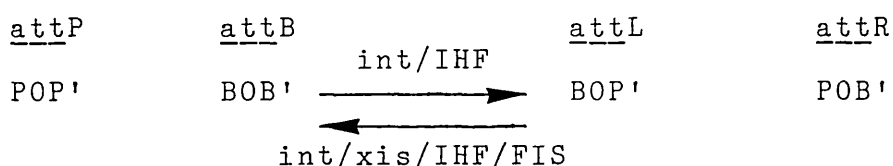
regulation of host range genes by inverting segments of DNA (Iida et al., 1982; Kamp et al., 1979; Plasterk et al., 1983; Zieg et al., 1977). All of these DNA invertases are able to complement each other (Iida et al., 1982). The establishment of persistent infection of Herpes Simplex Virus I in humans (Chou and Roizman, 1985), mating type switches in yeast (Nasmyth and Tatchell, 1980), the differential expression of a large repertoire of variant surface glycoproteins (Hoeijmakers et al., 1980) and switching of constant regions of immunoglobulin genes in mammals (Early and Hood, 1981) also involve site-specific recombination.

1.7.1 Site-specific Integration of Bacteriophages into the Genome of Bacteria:

Studies on the properties of integration of many bacteriophages including lambda, P4 and P1 have suggested the involvement of a site at the ends of the phages (attP), a sequence at the insertion site (attB) and a site-specific recombinase. In the case of lambda integration and excision the involvement of host proteins has also been established. For site-specific recombination of P4, P1 (a low frequency event) and lambda, two sites are required which are PattP + PattB, loxP + loxB and attP, + attB respectively. P4 and P1 which can also replicate as plasmids in the cell, frequently utilise this system to segregate the dimers into monomers (Abremski et al., 1983; Deho et al., 1984).

The overall steps involved in site-specific integration and excision of P4 and P1 are analogous to phage lambda. However the facts for lambda are based on detailed in vivo and in vitro studies some of which are discussed below.

Phage lambda integrates into the chromosome of E. coli by a regulated site-specific recombination event which requires two proteins, the lambda int and integration host factor (IHF) protein (Nash, 1981). Three proteins are required for excision xis, IHF protein and an additional small protein (12 Kdal) called "factor for inversion stimulation" (FIS) which has been shown to bind to the specific region in the lambda attP (Thompson et al., 1987). The overall recombination reaction can be summarised as follows;



In this formulation P, P', B and B' are distinct segments of phage (P) and host (B) DNA and "O" the crossing-over point of sequence in both att sites. The phage (attP) and host (attB) recombine to yield left (attL) and right(attR) sites in the inserted phage. The breaking and rejoining of DNA strands (phage and bacteria) is catalysed by integrase (a site-specific topoisomerase) in a recognition sequence (5'-CAATTNNT-3') at which the enzyme acts by making a 7 bp

staggered cut in the crossing-over region (Craig and Nash, 1984).

There are seven int binding sites in attP (2 in P, 2 in O and 3 in P') which is approximately 230 bp in length and three int binding sites in the attB which is 25 bp (Echols, 1986; Nash, 1981).

Electronmicroscopic and topological studies indicate that attP enters in the recombination reaction as a part of a folded nucleoprotein structure called an "intasome". It has been shown that the int protein condenses the 230 bp (800Å) of attP (POP') DNA into a tight complex about 140Å in diameter. The int binding sites are not contiguous or in the same orientation, therefore the DNA is folded or wound (Echols, 1986).

Bacteriophage P4 integration occurs in a 20 bp core sequence located at the 3' end of a leucine tRNA gene in the E. coli chromosome. P4 attP is flanked by a 5 bp repeated sequence which is structurally similar to the lambda attP region (Pierson and Kahn, 1987), however P4 integrase causes a 2 bp staggered cut at the core sequence (Argos et al., 1986). The integrases of seven different bacteriophages; lambda, ϕ 80, P22, P2, 186, P4 and P1 may use the same amino acid residues for stable integration of the phages in the host cell as they share a reasonable amino acid homology (Argos et al., 1986). However they cause different base pair staggered cuts at their recombining

sites and the recognition sites in each case are different.

In well characterised prokaryotic site-specific systems, the recombinases introduce short, staggered nicks within their homologous cross-over sequences during the in vitro recombination reaction. For example, phage lambda int protein, phage P1 cre protein and Tn3/gammadelta tnpR resolvase generate 7, 6 and 2 bp staggered cuts respectively. At the point of cleavage, int and cre become covalently attached to the 3' phosphate of DNA and produce a free 5' hydroxyl end, whereas tnpR resolvase connects to the 5' end of DNA and leaves the 3' hydroxyl terminus (Argos et al., 1986; Grindley and Reed, 1985).

1.7.2 Cointegrate Resolution; Another Site-Specific Recombination System:

The transposition among the members of the Tn3 family of transposons proceeds through an intermediate cointegrate molecule which is then the substrate for site-specific recombinase (resolvase) to perform a strand exchange reaction at synapsed res sites within the two elements. The resolution of a model cointegrate has been achieved in vitro which requires only supercoiled DNA containing the ²⁺ direct copies of the res sites, resolvase protein and Mg (Krasnow and Cozzarelli, 1983). The primary role of supercoiling is to facilitate the formation of a wrapped, functionally active structure like the intasome.

The res sites of gammadelta and Tn3 are approximately 120 bp and contain three separate resolvase binding sites as shown in figure 1.9. Each site consists of an imperfect inverted repeat of a 9 bp segment with a consensus sequence (5'-TGTCYNNTA-3') and is separated by a sequence of variable length (10 bp in site I, 16 bp in site II and 7 bp in site III). The cross-over point for recombination is in the centre of site I in the palindromic sequence (5'-TTATAA-3'). All these sites are required for efficient recombination and probably 12 resolvase monomers (6 dimers) occupy the 6 res sites present on two elements present in the cointegrate molecule (Kitts et al., 1983).

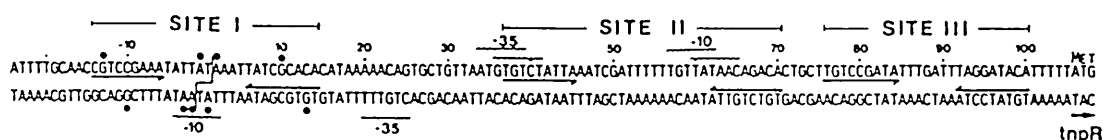


Figure 1.9 The res site of gammadelta. Sites I, II and III are binding sites for resolvase protein. The cross-over point is marked in Site I. The horizontal half arrows show the binding sequences by the C-terminal domain of resolvase. Stars mark the positions of the G and A nucleotides within Site I which inhibit cointegrate resolution when methylated in vitro (see text). The -10 and -35 regions of tnpA and tnpR promoters are also shown (taken from Grindley and Reed, 1985).

Flavey and Grindley (1987) carried out methylation and ethylation experiments on the res site of the gammadelta transposon. They found that methylation within a 9 bp consensus sequence inhibits resolvase binding to that site and ethylation of certain phosphates within and adjacent to this sequence also inhibits binding. They suggested that inhibition of resolution by methylation may be due to minor groove contact near the cross-over site.

Gammadelta resolvase is the most extensively studied recombinase. It consists of 183 amino acids and has two structurally and functionally distinct domains; 140 amino acid N-terminal and 54 amino acid C-terminal site (Abdel-Meguid et al., 1984). The smaller fragment has DNA binding specificity and the large fragment mediates protein-protein interaction, a step necessary for DNA strand exchange between the two res sites which are located on the cointegrate molecules. The crystallographic studies indicate that the two domains are joined together by a flexible hinge (Abdel-Meguid et al., 1984). This may account for the ability of resolvase dimer to bind cooperatively to individual binding sites with different geometric configuration.

A number of plasmid mediated site-specific recombination systems in prokaryotes have been identified which play a role in plasmid segregation during cell division (Dodd and Bennett, 1987; Summers and Sherratt, 1984). These plasmids

display homology with the tnpR gene product and the res site of the transposon Tn3 and gammadelta (Dodd and Bennett, 1987) which suggests that they may have been derived from Tn3-like elements.

It has been suggested in the 2 μ plasmid of Saccharomyces cerevisiae that intra-molecular site-specific recombination events between two inverted repeats of a 590 bp segment also mediate the resolution of plasmid multimers (Volkert and Broack, 1986).

1.7.3 Site-specific Inversion of DNA:

This class of recombination is responsible for the inversion of H flagellin genes of Salmonella, G segment of bacteriophage Mu, C segment of phages P1 and P7, and P element of e14 (a defective prophage in E. coli) by the gene products of hin, gin, cin and pin respectively.

All four recombinases of this class are able to complement each other (Iida et al., 1982; Kamp and Kahmann, 1981) and have more than 60% homology at the amino acid level. Comparison of the cross-over sites for these inversions revealed a 26 bp consensus sequence (5'-TTATCAAAACCT■TGGTTTAAGAGAA-3'), consisting of a virtually identical 12 bp inverted repeat with the central dinucleotide TT or (TC for G inversion), where (■) is the centre of symmetry at which strand exchange takes place (Hiestand-Nauer and Iida, 1983; Kahmann et al., 1985;

Plasterk et al., 1983; Plasterk and van de Putte, 1985; Zieg and Simon, 1980).

1.7.4 Requirements for Site-specific Recombination:

In vitro studies involving lambda integration into the chromosome, Tn3 cointegrate resolution and the inversion of DNA segments, have shown that there is a variable degree of complexity in the requirements for site-specific recombination. This variation is mainly based on the requirement for recombinational enhancer sequences, ATP and host factor proteins.

In vitro studies indicate that a 60 bp sequence located within the N-terminal part of the coding region of Hin gene stimulates inversion of the H segment 150 fold and can work in any orientation at numerous locations (Johnson and Simon, 1985). A similar situation was observed for G inversion that a 60 bp cis-acting sequence termed cix, stimulates the inversion about 15 fold in an orientation independent manner (Kahmann et al., 1985). The orientation and position independent nature of these cis-acting elements is reminiscent of transcriptional enhancers. It has been suggested that cis-acting elements may have protein binding sites involved in a reactive nucleoprotein structure. The common theory for transcriptional enhancers is that they provide the bidirectional entry site for polymerase II or one of its subunits (Khoury and Gruss,

1983). This mechanism may also apply to the entry of site-specific recombinases in the process of site-specific recombination.

The strand exchange mechanisms in the case of site-specific recombination differ on the basis of their requirement for high energy cofactors. In the case of gammadelta resolvase it has been shown that energy is conserved in a phosphoprotein linkage during strand exchange and religation which involves a serine residue (Reed and Moser, 1984). Therefore, there is no energy requirement for gammadelta cointegrate resolution. In several topoisomerases and the yeast FLP protein, a tyrosine residue has been identified as an active amino acid (Cozzarelli et al., 1984). In both cases the recombinational potential is dependent on the formation of the correct synaptic complexes.

Host-encoded proteins which may interact with DNA sequences for the formation of a synaptic complex have been identified for lambda integration as IHF, for lambda excision as IHF and FIS, for G inversion as Hu and for H inversion as FIS and Hu. The Hu protein is a heterodimer and has been shown to promote the underwinding of DNA in the presence of topoisomerase I to form the nucleosome-like configuration and displays a significant homology with IHF (Miller, 1984). FIS is another protein which has been shown to stimulate the excision of lambda from the

chromosome (Thompson et al., 1987).

A possible involvement of these proteins may be to communicate between the DNA molecules. The assembly of the recombination sites forms into synaptic complexes in which the DNA strands are aligned properly so forming the prerequisite for crossing-over. If this assembly is the rate limiting step, then host proteins that bind to cis-acting elements may facilitate the formation of this complex and increase the rate of site-specific recombination. From many different studies of the communication between two domains of DNA, there seems to be no general mechanism. Ptashne (1986) has divided the mechanisms into four groups; looping, twisting, sliding and oozing.

The direct demonstration of the looping model is based on the studies of the lambda operator, in which the interaction between two operators has been shown to loop out the intervening DNA (Hochschild and Patashne, 1986). This interaction is responsible for stable maintenance of lambda lysogeny in E. coli chromosome. According to the idea of looping out, it is the interaction between two DNA bound proteins, not the looping per se that regulates gene expression. If the intervening DNA is not long then the operator binding sites must lie on the same side of the DNA to allow the bound repressor proteins to make contact, i.e. an integral number of helical turns apart. The addition of a half integral turn has been shown to cause a failure to

loop out DNA in vitro (Hochschild and Ptashne, 1986). Electronmicroscopic studies revealed that the looped out DNA does not contact the proteins. Other examples of looping out mechanisms involve the efficient suppression of gal operon of E. coli (Irani et al., 1983) and the efficient expression of early gene of simian virus 40 (SV40) (Takahashi et al., 1985).

Twisting involves unwinding and altering the configuration of the DNA by proteins which bind directly to some altered form of DNA. The role of Hu in site-specific recombination may be exerted by this mechanism.

Oozing has been shown not to be the mechanism used in Tn3 cointegrate resolution because the amount of Tn3 resolvase sufficient to catalyse the reaction is insufficient to cover the DNA between the two res sites (Krasnow et al., 1983). Mechanisms involving looping and oozing are unprecedented in the process of transposition.

Sliding was once the favoured model for the resolution of Tn3 cointegrates. Recognition of the res sites by resolvase protein, formation of the synapsis and eventually strand exchange are three steps in cointegrate resolution. Before 1984, most models proposed the one-dimensional diffusion of resolvase for proper alignment of the two res sites (Krasnow and Cozzarelli, 1983). The mechanism now appears to be a two step synapsis model which was proposed by Boocock et al. (1986) from studies of Tn3 resolvase in

vitro. This model has demonstrated how antiparallel sites II and III interact with each other followed by twisting. In this model inter-molecular recombination and intra-molecular inversion can also occur.

1.7.5 DNA Flexibility as a Determinant in Site-specific Recombination:

The formation of a nucleosome-like structure for lambda integration, Tn3 family cointegrate resolution and H flagellin inversion has suggested a requirement for flexibility in the structure of DNA. Therefore the bending of intervening DNA is essential to bring the two domains into close proximity. There seems to be no generalised rule for bendability of DNA.

The tight bending of the DNA in eukaryotic nucleosomes (Klug and Trevor, 1986) and in catabolite activator DNA-protein complex of E. coli (CAP complex) (Fried and Crothers, 1983) is correlated with the occurrence of particular short sequences. In the nucleosome the trinucleotide AAA or TTT is often found at the minor groove which points towards the histone octamer whereas another trinucleotide GGC or GCC is located on the outward-facing minor groove. This particular base stacking interaction accommodates the deformations on grooves. The magnitude of curvature in CAP complex is the same as in nucleosomes and is directed towards the protein. An observation was

made by Wu and Crothers (1984) who found a bending locus contained in a Sau3AI restriction fragment from Leishmania tarentolae kinetoplast DNA. They identified the sequence which has a regular repeat of the sequence element CA₅₋₆ with 10 bp periodicity.

An electrophoretically anomalous DNA fragment was shown to exhibit a 10 to 11 bp periodicity in the distribution of AA and TT dinucleotides (Koo et al., 1986). Using chemically synthesised oligonucleotides, the observation was made that a tandemly repeated decamer (5'-NAAAATTTTN-3') had anomalous migration in polyacrylamide gels in comparison to the (5'-NTTTTAAAAAN-3') (Ulanorsky and Trifonov, 1987). This suggests that DNA bending occurs at AT rather than TA dinucleotide.

To demonstrate the requirement of bendability versus intrinsic bends of DNA at site I within the res site, Hatfull et al. (1987) carried out experiments on the res site of gammadelta. They found that resolvase induces a structural change in the DNA at the recombination cross-over point in site I at 5'-TTA/TAA-3'. The slash in the sequence represents the centre of a bend in site I of the res site. Mutant proteins which have a single amino acid change at either residue 118 or 128 from glutamate to lysine, failed to induce bending of site I. The assay for the bending was based on the observation that the structural changes induced by the wild-type resolvase

result in enhanced intercalation with a foot-printing reagent methidiumpropyl-EDTA.Fe(II)(MPE.Fe(II)). The mutant proteins which were unable to induce a bend in the DNA could not promote the intercalation. These in vitro studies have clearly shown that a localised kink facilitates the site-specific recombination for cointegrate resolution. Therefore it can be assumed that in other biological systems where proteins are bound to two different DNA domains (both in prokaryotes and eukaryotes), there is often an intrinsic bend in the DNA making interaction possible. In the absence of such bendability in DNA, this interaction otherwise would be energetically unfavourable.

1.8 Eukaryotic Transposable Elements:

The presence of transposable elements was first evidenced many years ago in eukaryotes by McClintock (1951). In recent years much of the attention has been focussed on the structural and functional organisation of eukaryotic transposons with the great attraction of understanding the evolution of transposable elements. Ty elements of Saccharomyces cerevisiae (Fink et al., 1980); P and copia or copia-like elements of Drosophila melanogaster (Rubin et al., 1983); Tc element of nematode Caenorhabditis elegans (Emmons et al., 1983); Tam element of Antirrhinum majus and Ac-Ds receptor regulator pair of elements of maize (Doring and Starlinger, 1986) are all thoroughly investigated

examples.

In the genome of Saccharomyces cerevisiae, the family of transposons (Ty) is widely distributed. Ty element consists of 5.6 Kb of DNA and is flanked by an approximately 330 bp of a directly repeated sequence called delta. Ty element has been shown to cause a number of specific mutations in different genetic loci, for example, at His4, Adr3 and Cyc7 (Roeder and Fink, 1983). Ty element causes a 5 bp duplication at the target site. The transposition of Ty element has been shown to be regulated by mating type switching. Many transposition events have been reported in the 5' non-coding region of these genes. Another transposable element of yeast (Sigma) has been shown to insert at the 5' end of a number of yeast tRNA genes (Del Rey et al., 1982).

The transposition of Ty element was shown to occur via an RNA intermediate which is a characteristic of vertebrate retroviruses (Boeke et al., 1985). Garfinkel et al. (1985) were able to show virus-like particles in the yeast cell under galactose promoted induction of transposition of Ty element. These particles resembled the intra-cisternal A-type particles (IAP) of mouse and copia particles of Drosophila. The IAPs are non infectious virus-like particles found in the mouse oocyte and pre-implantation embryo, and can transpose to new locations (Georgiev et al., 1984).

In Drosophila melanogaster 10% to 15% of the genome consists of moderately repetitive DNA and about one half of this is made up of 30 to 40 families of transposable elements with very similar properties. Some of the best studied of these are P element, copia, 412, 297, mdg1, mgd3 and B104 families (Finnegan, 1985). Several mutations in the biothorax complexes have been reported to be associated with the insertion of copia-like element. P element of Drosophila has been shown to be the casual agent of P-M hybrid dysgeneis which results in high rates of sterility, mutation and DNA rearrangements. The transposition of P element (P-M dysgeneis) occurs at very high levels only when P elements are introduced into an M cytotype egg (Engles, 1983).

Sixteen independent transpositions of P element have been found in a consensus sequence 5'-GGCCAGAC-3' (which is duplicated after insertion) in the white gene locus of Drosophila melanogaster (O'Hare and Rubin, 1983). Another element "hobo" of D. melanogaster is shown to cause a 8 bp target duplication and 11 independent insertions were found in a consensus sequence of either (5'-NNNNNNNAC-3') or (5'-CTTTTNN-3') (Streck et al., 1986).

The copia element of Drosophila melanogaster resembles the F family of Drosophila (Finnegan et al., 1978), the TED element of Trihoplausia ni (Miller and Miller, 1982) and retroviral provirus of vertebrates (Varmus, 1982). The

common feature of these elements is their ability to insert through an RNA intermediate.

Temin (1980) proposed an evolutionary relationship between the retroviruses (RNA tumour virus) and the eukaryotic transposable elements. It has been suggested in eukaryotes, that highly dispersed repetitive sequences (Alu family or fold back families) are thought to be the cause of the spread of transposable elements throughout the genome involving reverse transcription (Georgiev et al., 1984). Therefore it can be assumed that in eukaryotes those elements which transpose via reverse transcription could be evolved from retroviral provirus, Hepatis B-virus or Caulimosaic virus, as the replication of these virus particles proceeds through an RNA intermediate (Temin, 1985).

1.9 Evolution of Prokaryotic Transposable Elements:

Sapienza and Doolittle (1980) have proposed that transposable elements confer no selective advantage upon the organism that they inhabit, therefore they could be considered as selfish DNA. However this may not be true in all cases, for example, IS50 has been shown to improve the growth rate of E. coli without transposition (Hartl et al., 1983).

IS elements isolated from many different sources display many differences in their organisation (Iida et al., 1983). This may indicate a constraint on them for maintaining

their transpositional function. Changes could inactivate the transposition functions or alter them, or could result in increased transposition frequencies. In the latter case it is an advantage for the element over other DNA sequences. However a control is required to ensure a balance between the high frequency of transposition and the effect of such movements on the cell.

The composite transposons are thought to have evolved recently as the antibiotic resistance is not generally associated with the transposition function of the element. Therefore evolution of transposable elements may involve two steps. The best examples are Tn5 and Tn10 where the transposase encoded by the left IS of transposons has been inactivated by mutation(s); in the case of Tn5 of a single bp, and in the case of Tn10 of about 30 bp. These changes reflect the age of the transposon as Tn10 looks more ancient than Tn5. The ability of transposable elements to rearrange the DNA sequence makes them a major force in evolution. Transposable elements are not only causative agent for spreading antibiotic resistance, but they also carry operons from one organism to another. Amongst the Enterobacteriaceae, E. coli is the only species to have the lac operon. It is proposed that E. coli has acquired the lac operon from another bacterial species via a transposon Tn951 which contains the lacI, lacZ and lacY genes of E. coli (Cornelius et al., 1978).

The evolution of class II elements could be best demonstrated by comparing the origin of the transposition process, direction of transcription of tnpA, tnpR genes and the position of the res site which suggest that the ancestor transposon may have a tnpR gene on an invertible element (Schmitt et al., 1985). Gene products of tnpR gene and the res site of Tn3 family transposons have strong homology with the tnpR gene product and per site of plasmid R46 and are interchangeable (Dodd and Bennett, 1987). This suggests an evolutionary link between plasmids and transposons in which case transposable elements may have taken the cointegrate resolution function from the plasmid.

By focussing on sequences necessary for high frequency site-specific transposition (hot spot utilisation), one can gain insight on how target DNA is recognised by transposition machinery. The evolution of site-specific transposable elements (for example, Tn554, Tn1545, ISH1 and Tn7) can not be completely foreseen unless the contribution made from the target site can be found which ensures their insertion.

In this study availability of the target site for efficient transposition of the bacterial transposon Tn7 in the chromosome is examined. This implies another determinant for the spread of the transposable element among different bacteria and may focus on an evolutionary significance for the transposable element or target site.

1.10 Tn7 Transposition and Requirements:

Among the bacterial transposons, Tn7 has a distinct personality. It is an unusually large (14 Kb) transposable element and encodes resistance to the antibiotics trimethoprim, spectinomycin and streptomycin.

An interesting feature of Tn7 is its trimethoprim resistance determinant. Trimethoprim, like methotrexate and aminopterin, is a folic acid analogue which competitively inhibits the enzyme dihydrofolate reductase (DHFR). This enzyme is vital for a number of metabolic pathways including DNA synthesis. Tn7 encoded DHFR shows a greatly reduced affinity for trimethoprim. The complete nucleotide sequence of the trimethoprim resistant dihydrofolate reductase gene of Tn7 (with a deduced polypeptide of 157 amino acid residues) shows a significant homology with the E. coli and Lactobacillus casei chromosomal encoded DHFR enzymes. A change at position 30 from aspartate to glutamate in one out of the five conserved trimethoprim binding residues may be involved in the altered affinity to trimethoprim (Fling and Richards, 1983)

In recent years due to selection by clinical use of trimethoprim, many bacterial strains carrying incM plasmids that include Tn7 have been identified in hospital isolates (Datta et al., 1980). Tn7 has been specifically implicated in the dissemination of resistance by stably integrating into the chromosome of many pathogenic Enterobacteria (Towner et al., 1982). Another transposon Tn402 also

encodes trimethoprim resistance, but its sequence has not been determined. The comparison of the polypeptide predicted by the aminoglycoside-modifying enzyme 3' (9) -O-nucleotidyltransferase (aad) gene of Tn7 which mediates resistance to spectinomycin and streptomycin and that of Tn554, indicates that the genes are related (Fling et al., 1985), while the comparison of DNA sequence of Tn7 and plasmid R538-1 indicates that their aad genes are nearly identical.

Tn7 was first isolated from a trimethoprim resistant strain on the plasmid R483 (Hedges et al., 1972). Later many plasmids were shown to have insertions of Tn7 (Barth et al., 1978). Forty-nine independent insertions of Tn7 in the RP4 were used to identify 19 tra mutants in conjugal systems of RP4 (Barth and Grinter, 1977). Several random Tn7 insertions in the octopine utilisation gene of the Ti plasmid of Agrobacterium tumefaciens (Hernalsteens et al., 1978), the transfer region of plasmid R91-5 of Pseudomonas aeruginosa (Moore and Krishnapillai, 1982) and oriV region of plasmids R18 and R68 (Nash and Krishnapillai, 1987) have been isolated.

Tn7 transposes into most plasmids with a low frequency of $1E-6$ per cell per generation (which is similar to the transposition frequencies of many transposons) and at more or less random sites. An exception is the megaplasmid of Rhizobium meliloti which has a preferred site of

insertion (Bolton et al., 1984). The Tn7 insertion site of this plasmid is located outside the nif and nod gene regions. In the cases reported, Tn7 inserts into plasmids in only one of two possible orientations with respect to a fixed plasmid reference.

The most fascinating aspect of Tn7 is its transpositional properties to the chromosome of enteric bacteria in which Tn7 transposes at a very high frequency of $1E-2$ per cell per generation (Barth et al., 1978). All characterised transpositions of Tn7 into the E. coli chromosome are found at a single site at minute 82 and in a unique orientation (Lichtenstein and Brenner, 1981). In Pseudomonas aeruginosa, Tn7 inserts with high frequency at a specific site (Caruso and Shapiro, 1982). Nine independent insertions of Tn7 in the chromosome of Caulobacter crescentus were found at a single site (Ely, 1982). Tn7 has also been shown to insert preferentially at a site in the ilv gene cluster on the chromosome of vibrio species (Thomson et al., 1981). Its specific and efficient insertion into the chromosome of Xanthomonas campestris pv campestris was also reported (Turner et al., 1984).

When a region of the E. coli chromosome containing the preferred site is cloned into a plasmid, the high frequency and specificity of Tn7 transposition is retained (Lichtenstein and Brenner, 1982). The transposition frequency of Tn7 from the episomal element to this cloned target site was estimated to be $1E-3$ per cell per

generation (Lichtenstein and Brenner, 1982; Rogers et al., 1986). Recent evidence suggests that an insertion sequence element from Pseudomonas aeruginosa, IS21 seems to give a 400 fold increase in the transposition frequency of Tn7 onto plasmid R68.45 (Schurter and Holloway, 1987). This may be due to some complementation between the trans-acting transposition function of the two elements.

As with many other transposons including Tn3, gammadelta, phage Mu and IS2, the transposition of Tn7 causes 5 bp duplication at the insertion site. However, there is no direct evidence for a similar (replicative) mode of transposition for Tn7 like these elements. In contrast, different lines of evidence suggest that Tn7 transposition may involve a conservative mode of transposition. It was observed that deletion of any transposition gene does not lead to cointegrate formation which suggests that Tn7 does not encode a specialised resolution system (Hauer and Shapiro, 1984; Smith and Jones, 1984). However, Hodges (1983) observed in vivo cointegrates of Tn7 transposition and these were later found to be the result of recA independent homologous recombination rather than transposition (David Sherratt, personal communication). On the basis of latest information on Tn7, it seems that products of Tn7 transposition are simple insertions. The evidence for the multiple insertion of Tn7 in different plasmids (Barth and Grinter, 1977; Barth et al., 1978; Moore and Krishnapillai, 1982) suggests that Tn7 does not

encode a resolution system. Whether Tn7 produces simple insertion by conservative mechanism (i.e. cut and paste) or by replicative mechanism (i.e. a copy is generated before DNA translocation) has to be established.

Tn7 has unusual asymmetrical termini; both ends show close homology (75%), however the extreme ends (about 30 bp) shows near perfect inverted repeats. Both ends of Tn7 contain several copies of a 22 bp recurring motif. There are 5 direct repeats in the right end of Tn7 in continuous fashion, and 3 in the left end in discontinuous fashion. In the left end the repeats are separated by the sequence 5' GCTTTTATACTAA-3' which shows 10 out of 15 bp homology with the core region of bacteriophage lambda when allowance is made for gaps in some nucleotide positions (Lichtenstein and Brenner, 1982).

A detailed analysis of the att site from the E. coli chromosome has shown that Tn7 inserts in a region of dyad symmetry that serves as the rho independent transcriptional terminator of the glucosamine synthetase gene, glmS (Gay et al., 1986). The glmS gene encodes glutamine, amido-transferase and glucosamine 6-phosphate synthetase which are essential for the growth of the bacterium (Walker et al., 1984). The DNA sequence analysis of a 958 bp DNA fragment encompassing the attTn7 site suggests that the glmS coding region terminates at position 61 from the right end of the attTn7, and downstream there is a region of dyad symmetry followed by a run of T

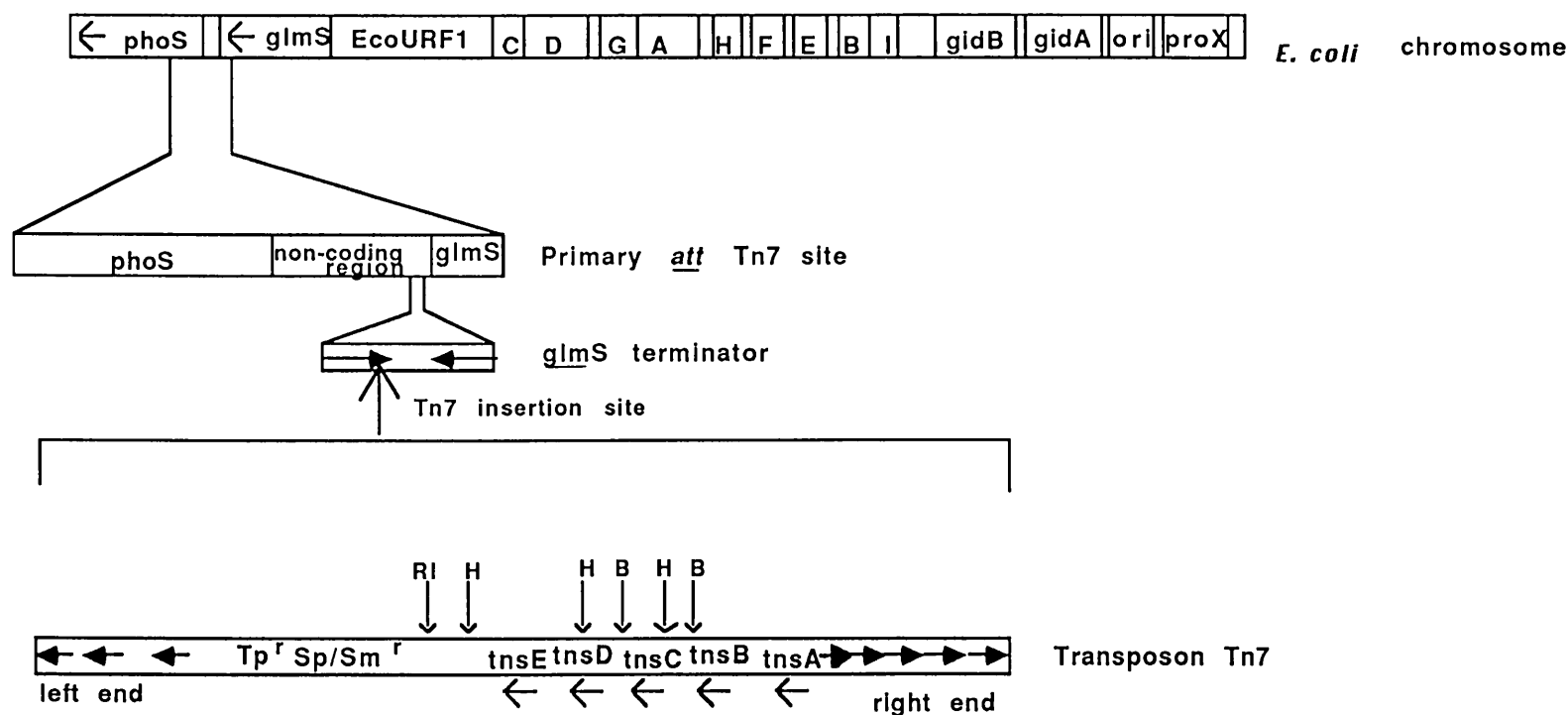


Figure 1.10 Genetic map of the *attTn7* site from *E. coli* chromosome. (Walker *et al.*, 1984; Surin *et al.*, 1984). The 958 bp *EcoRI* RE fragment serves as a target for Tn7 transposition when cloned into plasmids (Lichtenstein and Brenner, 1982). The insertion of Tn7 occurs at position 86 in the *glmS* terminator starting from the *glmS* end (right end) of *attTn7*.

KEY

- ← inverted repeat
- ← direction of translation
- R I EcoRI
- H HindIII
- B BamHI

residues which forms the glmS transcriptional terminator with a very high thermodynamic value ($\Delta G = -21.2$ Kcal per mol) into which Tn7 transposes (Gay et al., 1986).

The phoS gene starts at position 373 in the attTn7 DNA, and between positions 61 to 372 is the non-coding region (see figure 1.10). The Tn7 insertion point lies within this fragment at position 86 from the glmS (right end) end in the transcriptional terminator of the glmS gene which has potential to form a cruciform structure. The DNA sequence at the site of insertion in plasmid ColE1 also shows potential for a secondary structure (Lichtenstein and Brenner, 1982).

S1 mapping studies indicated that the presence of Tn7 in the chromosome of E. coli abolishes termination of glmS transcription at this site and transcription was found to extend into the right end of Tn7. These transcripts terminate at a new site "Tm", 127 base pairs from the right end of Tn7 (Gay et al., 1986).

Gay postulates that transcriptional read-through would occlude transcription from a site within Tn7 that has homology to the E. coli consensus promoter sequence and which is followed by an open reading frame (ORF) of a putative Tn7 transposase shown to have homology to the transposases of Tn3 and IS2 (Gay et al., 1986).

Since transcription from the glmS promoter is likely to be growth rate dependent then the rate of transposition should

be directly related to the level of transposase expression. Tn7 transposition from the chromosome site could be regulated by the growth rate of the cell. According to Gay's theory this inhibition of Tn7 transposase expression is advantageous for the optimum survival of the cell as this will not allow Tn7 transposition to secondary, possibly lethal, sites during the growth of the bacterium (Gay et al., 1986). This inhibition phenomenon could not be explained in transposition from plasmids to other sites.

The remarkable property of cis-acting transpositional immunity whilst shown by Tn3 and phage Mu, is not clearly defined in Tn7. Hassan and Brevet (1983) demonstrated in two steps that two copies of Tn7 (one with an additional antibiotic marker) can be inserted in plasmid RP4. This observation, in conjunction with multiple insertions of Tn7 in different plasmids, provides a clue that a replicon carrying a copy of Tn7 is not immune to further insertion. However, Hauer and Shapiro (1984) believe that Tn7 displays transpositional immunity. Another observation in favour for Tn7 transpositional immunity was made that when tandem copies of the E. coli attTn7 site are cloned into plasmid, Tn7 inserts into only one site (David Sherratt, personal communication). The apparent difference in these analyses may be due to different conditions used to measure the transposition frequency and therefore indicate a variable degree of Tn7 transpositional immunity.

A detailed restriction map of Tn7 was established by Gosti-Testu et al. (1983) (see figure 1.11). This shows that there is a unique EcoRI RE site in Tn7 which separates the transposition genes and the antibiotic resistance gene into 8.9 Kb and 5.9 Kb discrete fragments respectively.

A number of conflicting reports are published that define the regions of DNA required for Tn7 transposition (Smith and Jones, 1984; Hauer and Shapiro, 1984; Ouartsi et al., 1985). However, in all studies the 8.9 Kb EcoRI RE fragment to the right end of the transposon was shown to complement transposition of deleted derivatives of Tn7.

In a more precise study, the transposition function of Tn7 was mapped by complementing a "minimum" Tn7 which is deleted for all but the end of the transposon. This was complemented by different regions of Tn7 cloned into multi-copy plasmids. Using this genetic complementation test, five gene products required for Tn7 transposition were identified (Rogers et al., 1986) (see figure 1.11). Four of these (tnsA-tnsD) are required for insertion into the attachment site (the target used was 958 bp attTn7 from the E. coli chromosome) and the fifth tnsE, plus tnsA-tnsC are required for transposition into plasmid which lacks the preferred "hot site". It has been shown by gel retardation assays that tnsD specifically binds with the attTn7 (David Sherratt, personal communication).

Two promoters (P1 and P2) were identified in the right end

of Tn7 which indicates that at least tnsA, tnsB and tnsD are translated in the same direction, while expression from P1 appears to be repressed by tnsB (Rogers et al., 1986; Carlos Flores, personal communication). Four gene products encoded within Tn7 were identified using maxi-cell analysis by Brevet et al. (1985). The location of these proteins were in good agreement (except for tnsA) with the functional map determined by Rogers et al. (1986).

DNA sequence analysis of transposition (tns) region (8.9 Kb EcoRI fragment) of Tn7 has identified the 5 possible open reading frames for the 5 tns proteins (Carlos Flores, personal communication). The transcription of all genes are oriented in the same direction with the amino terminus of tnsA adjacent to the right end of Tn7 (see figure 1.11). The open reading frames of some transposition genes overlap with each other; this suggests a translational coupling between tns genes. For example, the tnsA overlaps with tnsB, tnsB with tnsC and tnsC with tnsD (see figure 1.11). However the fact that 5 proteins are encoded by Tn7 indicates that the transposition pathway of Tn7 is more complex in comparison to other elements. This complex genetic arrangement probably provides two alternative transposition pathways. The exact functions of these proteins in attTn7 and non-att target site selection has to be understood. It is possible that tns gene requirements for transposition are defined by the target of Tn7 transposition.

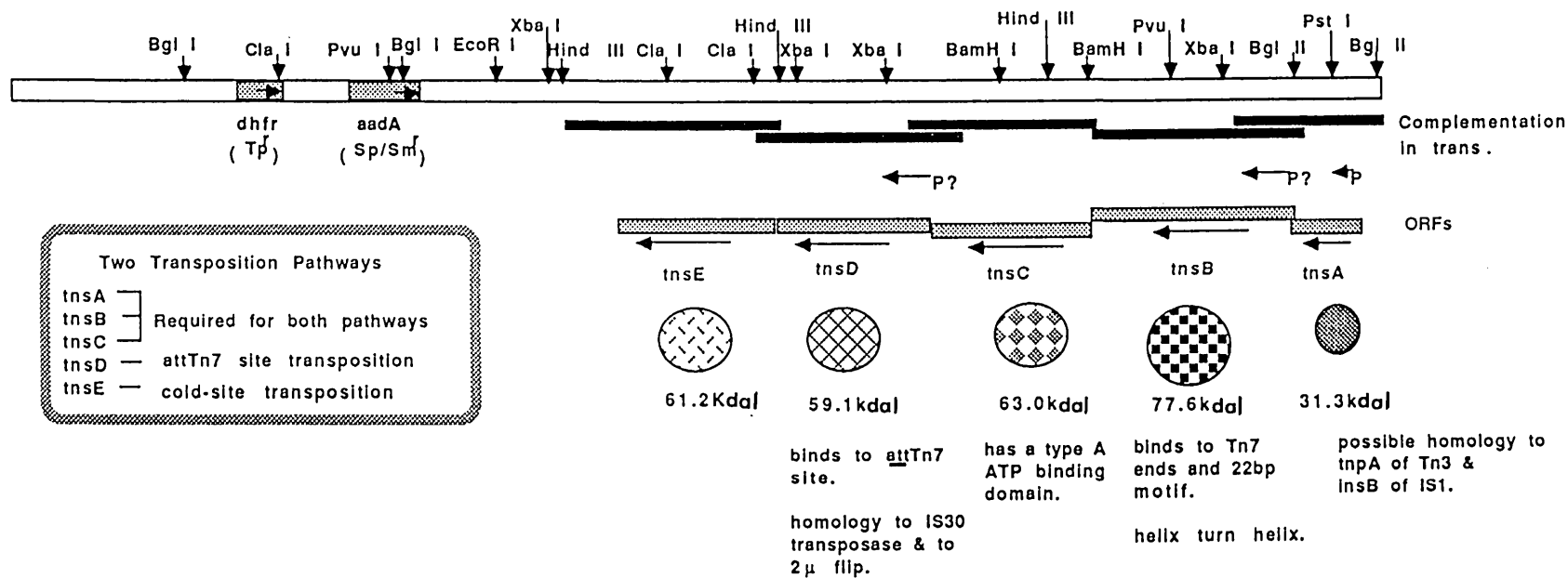


FIGURE 1.11 GENETIC ORGANISATION OF TRANSPOSON Tn7

1.11 Preliminary Account of the Project:

In order to understand the high frequency, site-specific transposition of Tn7 in the bacterial chromosome, several approaches were taken to define the requirement of DNA sequence at the attTn7 site which are the subject of this thesis.

Chapter 2 presents the materials and methods adopted for the work involved in this thesis.

Chapter 3 of the results section describes the development of a novel assay system to determine the frequency of transposition of Tn7 in vivo. Using this assay system, a 64 bp DNA sequence was identified by unidirectional deletion analysis of the E.coli 958 bp attTn7 site (chapter 4). A 70 bp fragment encompassing the mini attTn7 site was cloned in M13, mp10 which can function as an efficient and site-specific attTn7 site for Tn7 transposition.

Chapters 5 and 6 describe the strategies adapted to obtain the DNA sequence of attTn7 sites from the chromosomes of Agrobacterium tumefaciens, Klebsiella pneumoniae and Serratia marcescens for a comparison. The DNA sequence comparison of the attTn7 sites from different bacterial species and the deletion analysis of the attTn7 from the E. coli chromosome, led to the proposal of some models for involvement of different sequence elements within the attTn7 for high frequency transposition of Tn7.

The components of the attTn7 were analysed in a mini attTn7 site (70 bp DNA sequence) by constructing attTn7 sites with altered structures; the strategies of which are discussed in chapter 7.

Finally in chapter 8, results are presented on the effect of base substitution within a 70 bp attTn7. A large number of base substitutions were generated in the 70 bp attTn7 to determine the individual nucleotides responsible for high frequency transposition of Tn7. Three domains were identified within the 70 bp attTn7 by mutagenesis techniques which may be responsible for the formation of highly precise specialised nucleoprotein structure in which the DNA is probably folded or wound for site-specific transposition of Tn7.

Chapter 9 summarises the results of this work and concluding remarks are made with discussion on results.

CHAPTER 2

Materials and Methods

2.1 Materials:

2.1.1 Consumables:

Bacto-agar, tryptone, Mueller Hinton Medium (dehydrated), and yeast extract were purchased from Difco Laboratories, West Molesey, Surrey.

BSA, lysozyme, salmon sperm DNA, IPTG, X-gal, amino acids, Ficoll, polyvinylpyrrolidone, hydroquinone, HEPES, PMSF, chloramphenicol, trimethoprim, tetracycline, ampicillin, streptomycin, spectinomycin, gentamycin, kanamycin, and glucosamine were all purchased from Sigma Chemicals Ltd., Kingston-Upon-Thames, Surrey.

Ethidium bromide, DTT, caesium chloride, acrylamide, methylene-bisacrylamide (electrophoresis grade) and agarose "ultra-pure" were purchased from BDH Ltd., Poole, Dorset.

Urea "ultra-pure" was purchased from Swatrz-Mann, Orangeburg, N.Y.

Polyethylene glycol (6000) and dimethyldichloro silane solution were purchased from BDH.

Calf intestinal alkaline phosphatase, RNase A, DNase, DNA polymerase 1 (Klenow-fragment) and Avian Myoblastosis Virus (AMV) reverse transcriptase were purchased from Boehringer Corporation (London) Ltd., Lewes E., Sussex.

All isotopes were purchased from Amersham International, Amersham, Bucks.

Restriction endonuclease enzymes, M13 universal primer, hybridisation primer, exonuclease III, S1 nuclease were purchased from New England Biolabs (CP Labs., Bishops Stortford), Amersham and Bethesda Research Labs. (Cambridge).

Nitrocellulose was purchased from Schleiden and Schnell.

Phenol was purchased from Fissons and redistilled at Imperial College, London.

The mismatched heptadecamer was purchased from Professor Alan Fersht's laboratory at Imperial College, London.

pUC18 cloning site polylinker, Dextran sulphate (500,000) and Sephadex G50 were purchased from Pharmacia Fine Chemicals, Milton Keynes, Buckinghamshire.

2.1.2 Media:

AB media (10X) 30g K₂HPO₄ , 10g NaH₂PO₄ , 1g NH₄Cl, 1.5g
KCl, dissolved in 1 litre of

distilled water and adjusted pH to 7.4 with 5M NaOH. After autoclaving 25ml of 20% glucose, 0.6ml of 1M MgSO_4 , 0.09ml of 1M CaCl_2 and 0.09ml of 1M FeSO_4 were added separately.

L Broth 10g tryptone, 5g yeast extract, 5g NaCl, 1g glucose dissolved in 1 litre of distilled water and adjusted pH to 7.0 with 5M NaOH.

M9 (10X) 70g K_2HPO_4 , 20g KH_2PO_4 and 40g NH_4SO_4 dissolved in 1 litre of distilled water.

M9 media (1X) 0.1ml of 1M MgSO_4 , 12.5ml of 20% glucose and 0.1ml of 1M CaCl_2 were added separately in 1 litre of 1X M9 buffer.

MH media 38g of Mueller Hinton media were dissolved in 1 litre of distilled water and autoclaved.

2X YT 16g tryptone, 10g yeast extract, 10g NaCl dissolved in 1 litre of distilled water, adjusted pH to 7.2 with 5M NaOH and autoclaved.

Top Agar As for media, but 7.0g agarose or agar added per litre and autoclaved.

2.1.3 List of Bacterial Strains:

<u>E.coli</u> strains	Genotype	Source
C600	F',thi-1,thr-1,leuB6,lacY1, tonA21,supE44,lambda ⁻ .	Backmann (1972)
Δ7	As C600 but <u>glm</u> S ⁻ ,kan ^r .	C. Flores (Unpublished)
Δ7::Tn7 (1-7)	As Δ7 but with insertions of Tn7 in the chromosome. Seven independent events.	This study
DH5	F',endA1,hsdR17,(rk ⁻ ,mk ⁻), supE44,thi-1,lambda ⁻ ,recA1, gyrA96,relA (lacZYA-argF), v19,φ80, lacZ ΔM15.	Bethesda Research Labs.
HB101	F',hsdS20(r ⁻ m ⁻ _{B B}),recA13, ara-14,proA2,lacY1,galk2, rspL20(Sm ^r),xyl-5,mtl-1, supE44,lambda ⁻ .	Boyer and Roulland-Dussoix (1969)
71/18	K12, lac ⁻ ,pro,suIII ⁺ ,thi- F', proA ⁺ B ⁺ ,lacI ^q , lacZ, ΔM15.	Messing <u>et al.</u> (1977)
BMH 71/18 mutL	As 71/18 but mutL::Tn10.	Kramer <u>et al.</u> (1984)

71/18::Tn7	As 71/18 but with insertion of Tn7 in the chromosome.	This study
TG2	⁻ lac ⁻ ,pro ⁻ ,sup ⁺ E ⁻ ,thi ⁻ ,recA ⁻ , Srl::Tn10 (Tc ^r),hsd 5, EcoK ⁻ r ⁻ m ⁻ ,F,tra,36,proAB, lacI ^q ,lacZ ΔM15.	Gift from D. Glover, Imperial College
TG2::Tn7	As TG2, but with insertion of Tn7 in the chromosome of TG2.	This study
LP144	nusA1,galkK2,recA, rpsL,strR.	Gift from R. Hayward
LP145	galkK2,recA,srl::Tn10, rpsL,strR.	Gift from R. Hayward
LP144::Tn7	As LP144 but with insertion of Tn7 in the chromosome.	This study
LP145::Tn7	As LP145 but with insertion of Tn7 in the chromosome.	This study
Other bacterial strains	Genotype	Source
A136 <u>Agrobacterium tumefaciens</u>	C58 cured,rif(5),nalS.	Gift from E. Nester
OK8 <u>Klebsiella pneumoniae</u>	Prototroph.	Gift from E. Nester

MW62 <u>Serratia</u> <u>marcescens</u>	DNase ⁻ , prototroph.	Gift from R. Roberts
PAS279 RP4::Tn7 <u>Pseudomonas</u> <u>aeruginosa</u>	⁻ thr , ^r Kn , ^r Ap , ^r Tc , ^r Tp , ^r Sp , ^r Sm .	Fennwald and Shapiro (1979)
ATQ::Tn7 (1-10)	As A136 but with insertion of Tn7 in the chromosome. Ten independent events.	This study
KPQ::Tn7 (1-10)	As OK8 but with insertion of Tn7 in the chromosome. Ten independent events.	This study
SMQ::Tn7 (1-10)	As MW62 but with insertion of Tn7 in the chromosome. Ten independent events.	This study

2.1.4 List of Plasmids:

Plasmids	Description	Source
pCPL6	16 Kb fragment containing <u>att</u> Tn7 site cloned into ^r pBR322, Ap .	C. Lichtenstein (1980)
pCPL6::Tn7	Insertion of Tn7 in pCPL6, ^r Ap , ^r Tp , ^r Sm , ^r Sp .	C. Lichtenstein (1980)

pEAL1	A 958 bp fragment containing <u>att</u> Tn7 cloned into <u>EcoRI</u> RE site of pACYC184, Tc ^r .	C. Lichtenstein (1980)
ColE1	6.1 Kb colicin E1 producing plasmid.	Tyler and Sherratt (1975)
ColE1::Tn7	Insertion of Tn7 in <u>HaeII</u> RE fragment of ColE1, Tp ^r , Sm ^r , Sp ^r .	Gift from D. Sherratt
pRS5	Deletion of large <u>EcoRI</u> RE fragment from ColE1::Tn7, does not produce colicin.	C. Lichtenstein (1980)
RP4	49 Kb incP plasmid, Ap ^r , Tc ^r , Kn ^r .	P. Barth <u>et al.</u> (1978)
RP4::Tn7	Insertion of Tn7 in RP4, Tc ^r , Ap ^r , Kn ^r , Tp ^r , Sm ^r , Sp ^r .	Gift from D. Ribbons
pPH1JI	54.7 Kb incP plasmid, Gm ^r , Cm ^r , St ^r , Sp ^r .	P. Hirsch (1978)
pTJS140	7.7 Kb plasmid with ColE1 and RP4 type origin of replication, Ap ^r .	Gift from P. Zambryski
pTJS75Kn	6.5 Kb incP derived plasmid, Kn ^r .	Gift from S. Stachel
pUC18	2.7 Kb plasmid, Ap ^r .	Bethesda Research Labs.

pGC2	2.7 Kb plasmid with cloning site polylinker and a 250 bp GC rich DNA sequence (GC clamp).	Myers <u>et al.</u> (1985)
pGC270	70 bp <u>attTn7</u> site cloned into <u>Bam</u> HI, <u>Eco</u> RI RE sites of pGC2.	This study
pAIQ:: Δ Tn7 (1-2)	Approximately 6-8 Kb insert of Tn7 and <u>attTn7</u> from <u>A.tumefaciens</u> cloned into <u>Sal</u> I RE site of pTJS75Kn, r r Kn , Tp . Two events.	This study
pKIQ::Tn7 (1-6)	21 Kb insert of Tn7 and <u>attTn7</u> from <u>K. pneumoniae</u> cloned into <u>Sal</u> I RE site of r r r r pTJS75Kn,Kn ,Tp ,Sm ,Sp . Six events.	This study
pSIQI::Tn7	25 Kb insert of Tn7 and <u>attTn7</u> (first site) from <u>S. marcescens</u> cloned into <u>Sal</u> I RE site of pTJS75Kn, r r r r Kn , Tp ,Sm ,Sp .	This study
pSIQII::Tn7	20 Kb insert of Tn7 and <u>attTn7</u> (second site) from r r r r <u>S. marcescens</u> ,Kn ,Tp ,Sm ,Sp .	This study

2.1.5 M13 Derived Clones:

Name	Description	Source
M13, mp10	6.4 Kb single-stranded, filamentous, male-specific <u>E.coli</u> phage containing a cloning site polylinker in a <u>HindIII</u> RE fragment of lac regulatory region.	Messing (1983)
M13, mp11	As M13, mp10 but cloning site polylinker in reverse orientation.	Messing (1983)
VHV12	958 bp <u>E.coli</u> <u>attTn7</u> fragment cloned into <u>EcoRI</u> RE site of M13, mp8 with the right end of <u>attTn7</u> adjacent to the priming site of M13, mp8.	C. Lichtenstein (1980)
VHV22	As VHV12 but cloned into M13, mp10 in the reverse orientation.	This study
(Δ L)n	(n) number of base pairs deleted from the left hand end (L) of the 958 bp <u>attTn7</u> .	This study

(Δ R)n	(n) number of base pairs deleted from the right hand end (R) of the 958 bp <u>att</u> Tn7 fragment.	"
MAS	70 bp <u>att</u> Tn7 cloned into <u>Sma</u> I RE site in M13, mp10.	"
MP11GC	250 bp GC rich region (GC clamp) cloned into M13, mp11.	"
MP11 <u>ATT</u> GC	250 bp of GC clamp and 70 bp <u>att</u> Tn7 cloned into M13, mp11.	"
MASC	A <u>Cla</u> I RE site generated 5 bp outside the termination codon of <u>glm</u> S in the 70 bp <u>att</u> Tn7 DNA (MAS).	"
MASC::Tn7 (1-3)	MASC with natural insertion of Tn7, Tp ^r . Three events.	"
MASCPL77	57 bp insert (containing 55 bp pUC18 polylinker plus 2 bp target duplication) cloned into the <u>Cla</u> I RE site of MASC.	"
MASCPL73A	53 bp insert cloned into MASC (obtained by deleting 4 bp from the <u>Kpn</u> 1 RE site of MASCPL77 with S1 nuclease).	"

MASCPL73B	53 bp insert in MASC, (obtained by deleting 4 bp from the <u>Bam</u> HI RE site of MASCPL77 with S1 nuclease).	"
MASCPL69	49 bp insert in MASC, (obtained by deleting 8 bp from the <u>Bam</u> HI RE site of MASCPL77 by S1 nuclease).	"
MASCPL32	12 bp insert in MASC, (obtained by deleting 45 bp from the <u>Kpn</u> I RE and <u>Sph</u> I RE fragment by digestion and then treated with S1 nuclease).	"
MASCPL43	<u>In vivo</u> 12 bp insert and 1 bp deletion obtained in MASCPL32 when propagated in bacterial strain 71/18 ColE1::Tn7.	"
Δ1LOOP	15 bp deletion in the transcriptional terminator of <u>glm</u> S in MP11 <u>ATTGC</u> .	"
Δ2LOOP	22 bp deletion in the transcriptional terminator of <u>glm</u> S in MASC.	"

Δ3LOOP	20 bp deletion in the transcriptional terminator of <u>glmS</u> in MASCPL.	"
LJA1	<u>Sau3AI</u> left junction (Tn7- <u>att</u> Tn7) fragment from pAIQΔ::Tn7 cloned into <u>Bam</u> HI RE of M13, mp10 in orientation 1. *	"
LJK1	<u>Sau3AI</u> left junction (Tn7- <u>att</u> Tn7) fragment from pKIQ::Tn7 cloned into M13, mp10 in orientation 1. *	"
LJK2	As LJK1 but in orientation 2. *	"
LJS2.I	<u>Sau3AI</u> left junction (Tn7- <u>att</u> Tn7) fragment from pSIQI::Tn7 cloned into M13, mp10 in orientation 2. *	"
LJS2.II	<u>Sau3AI</u> left junction (Tn7- <u>att</u> Tn7) fragment from pSIQII::Tn7 cloned into M13, mp10 in orientation 2. *	"
RJK1	<u>Sau3AI</u> right junction (Tn7- <u>att</u> Tn7) fragment from pKIQ::Tn7 cloned into M13, mp10 in orientation 1. *	"

RJK2	As LJK1 but in orientation 2.	"
	*	
RJS2.I	<u>Sau</u> 3AI right junction	"
	(Tn7- <u>att</u> Tn7) fragment from	
	pSIQI::Tn7 cloned into M13,	
	*	
	mp10 in orientation 2.	
RJS1.II	<u>Sau</u> 3AI right junction	"
	(Tn7- <u>att</u> Tn7) fragment from	
	pSIQII::Tn7 cloned into M13,	
	*	
	mp10 in orientation 1.	

* Orientation 1 is the following order;
M13 priming site - Tn7 - attTn7;
orientation 2 is the reverse order, viz;
M13 priming site - attTn7 - Tn7.

2.2 Methods:

This section of the chapter is divided into four parts dealing with the "microbiological", "biochemical", "base substitution mutagenesis" and "in vitro DNA-protein interaction" studies.

2.2.1 Microbiological Techniques:

2.2.1(i) Antibiotic Selection:

Throughout this work the standard antibiotic concentrations given in table 2.1 were used in both liquid and plate selections. The appropriate volume of the antibiotic stock solution was mixed with the molten agar medium solution which had been pre-cooled to 45^o C.

2.2.1(ii) Transformation of E.coli with plasmid DNA:

The general procedure for transforming bacteria with plasmid DNA was carried out according to the method of Kushner (1978) as described by Maniatis et al. (1982).

The method used for transformation of M13 phage vector replicative form (RF) or single-stranded DNA to F' bacteria for obtaining the plaques was modified from Mandel and Higa (1970).

Table 2.1 Antibiotic resistance levels for different bacterial genera.

Antibiotics	<u>Escherichia coli</u>	<u>Agrobacterium tumefaciens</u>	<u>Klebsiella pneumoniae</u>	<u>Serratia marcescens</u>	<u>Pseudomonas sp.</u>	Stock Solution
Ampicillin (Ap)	50 ug/ml	200 ug/ml	200 ug/ml	200 ug/ml	100 ug/ml	100 mg/ml (H ₂ O)
Chloramphenicol (Cm)	13 ug/ml	-	-	-	-	34 mg/ml (Ethanol)
Gentamycin (Gm)	20 ug/ml	100 ug/ml	100 ug/ml	50 ug/ml	-	20 mg/ml (Methanol)
Kanamycin (Kn)	50 ug/ml	200 ug/ml	200 ug/ml	200 ug/ml	-	50 mg/ml (H ₂ O)
Spectinomycin (Sp)	25 ug/ml	300 ug/ml	-	-	-	100 mg/ml (H ₂ O)
Tetracycline (Tc)	15 ug/ml	-	-	-	-	15 mg/ml (50% Ethanol)
Trimethoprim (Tp)	100 ug/ml	-	500 ug/ml	500 ug/ml	100 ug/ml	50 mg/ml (50% Ethanol)

0.5ml of an overnight bacterial culture was diluted in 10ml of 2X YT liquid media and shaken at 37 °C until the O.D. was 0.4-0.6 units. The cells were centrifuged at 4 °C for 5 minutes at 10000 r.p.m., then resuspended in 2ml of ice cold 0.1M MgCl₂ and 20mM MgSO₄ and placed on ice for 10-15 minutes. After recentrifugation as above, the cells were resuspended in 0.5ml ice cold 0.1M CaCl₂ and kept on ice for 1 hour. The cells were used immediately for optimal transformation efficiency. To 0.2ml of this suspension 100ng of DNA were added and the mixture incubated at 0 °C for 1-2 hours. The cells were heat shocked at 42 °C for 2 minutes and chilled on ice for a further few minutes. To the aliquots of transformed cells were added 0.2ml of an exponentially growing culture of TG2, 50µg/ml IPTG and 50µg/ml X-gal (final concentration) and 3ml of 2X YT top agar which had been melted previously and kept at 45 °C. The mixture was immediately poured onto 2 X YT plates and these plates incubated at 37 °C overnight to obtain M13 plaques.

2.2.1(iii) Assay for the Loss of Colicin E1 Production (i.e. Loss of ColE1 Plasmid):

Colonies to be checked for the loss of ColE1 plasmid (i.e. Colicin production) were tooth-picked in duplicate onto LB plates (together with positive and negative controls) and grown at 37 °C. One plate was inverted over a drop of

chloroform (to kill bacteria) for 2-3 minutes and then overlaid with a fresh culture of Colicin E1 sensitive cells in 3ml of top agar.

Following incubation at 37^o C, Colicin producing colonies generated a clear zone around themselves where the lawn of sensitive cells had failed to grow. Colonies which had lost ColE1 failed to generate a clear region.

2.2.1(iv) Bacterial Conjugation:

The method for bacterial mating was adapted from Miller (1972). Donor and recipient cells were grown separately in 2X YT media at 37^o C to an O.D. of 0.4 units. The cells were then serially diluted and plated on 2X YT plates to obtain isolated colonies. These colonies were picked in a 1:1 ratio and mixed on 2X YT plates. All matings were incubated at 37^o C except those involving Agrobacterium tumefaciens and Pseudomonas aeruginosa which were incubated at 30^o C. The exconjugants were selected by means of an auxotrophic marker or antibiotic resistance (see table 2.1).

2.2.1(v) Curing of Episome and Plasmid:

Acridine orange was used to cure the F'episome and the incP plasmid at a concentration of 100µg/ml for E.coli, and 200µg/ml for other bacterial species.

One bacterial colony was picked from a plate and inoculated into 2X YT media in the presence of acridine orange and incubated overnight at 37 °C (Agrobacterium at 30 °C). The overnight culture was repeatedly subcultured by inoculating into acridine orange. The culture was streaked onto 2X YT plates, incubated overnight and resultant colonies tested for the loss of plasmid encoded antibiotic markers.

Bacterial colonies to be tested for the loss of F' were tooth-picked onto 2X YT plates containing an overlay of ampicillin transducing M13 phages with very high m.o.i.'s. As F' is required for the formation of the sex pillus and consequently for M13 infection, those bacteria which had lost the F' were unable to grow on this phage lawn.

2.2.1(vi) M13 Transduction Assay of Tn7:

A single fresh bacterial colony of 71/18 Col E1::Tn7 was picked from a minimal glucose plate containing trimethoprim and grown in 2ml of 2X YT media in a 50ml bacterial culture tube at 37 °C to an O.D. of 0.2 units. Then a single, well separated, M13⁵⁵⁰ plaque was picked with a sterile pasteur pipette and blown into the 2ml of growing culture. The cells were shaken vigorously at 37 °C for 5 hours, spun down at 10000 r.p.m. for 2 minutes and the phage supernatant saved.

Serial dilutions of the phage supernatant were made in phage buffer (10mM Tris.HCl, pH 8.0; 10mM MgSO₄). 0.1ml of phage solution from 10⁻⁷, 10⁻⁸ and 10⁻⁹ dilutions were mixed with 0.2ml of a TG2 culture (at an O.D. of 0.8 units) and added to 3ml of molten top agar at 45 C. The mixture was immediately poured onto 2X YT plates, incubated at 37 C overnight and the plaques were counted to determine the total number of plaque forming units (P.F.U.) per millilitre of phage solution.

Similarly, phage dilutions were mixed with a TG2 culture as above and kept at 37 C for 10 minutes for phage adsorption. The mixture was immediately plated onto Mueller Hinton plates containing trimethoprim and tetracycline. 0.2ml of three controls (i.e. 71/18 COLE1::Tn7, TG2 and phage supernatant) were also plated separately on MH, Tc, Tp plates. The plates were incubated at 37 C for 24 hours and trimethoprim transduced colonies were counted for each dilution. The transduction frequency was deduced by dividing the total number of trimethoprim resistant colonies per ml of the phage solution by the total number of P.F.U. per ml of the same phage stock. In order to avoid multiplication of the transducing phage, only 10 transduction assays were performed at a time.

2.2.2 Biochemical Techniques:

This section is subdivided into different categories (2.2.2A to 2.2.2F) according to the generalised and specific techniques.

2.2.2A General Techniques:

2.2.2A(i) Extraction of DNA with Phenol and Phenol/chloroform Mixture:

An equal volume of neutralised phenol (redistilled and equilibrated with 1M Tris.HCl, pH 7.5) was mixed thoroughly with the DNA solution. The mixture was spun (10000 r.p.m.) for 5 minutes. The upper, aqueous layer was transferred to another tube and traces of phenol removed by extraction with an equal volume of a chloroform: isoamylalcohol (24:1) mixture. The DNA solution was mixed thoroughly, spun (10000 r.p.m.) for 5 minutes and the upper layer transferred to a new tube for DNA precipitation.

2.2.2A(ii) Precipitation of DNA with Ethanol:

To the DNA solution 0.3M (final concentration) sodium acetate (pH 4.8), or 2.5M (final concentration) ammonium acetate (pH 7.5), and 2.5 volumes of absolute ethanol were added. The solution was kept on a dry ice /isopropanol bath

(-70^o C) for 10 minutes and the DNA pelleted by centrifugation at 10000 r.p.m. for 15 minutes. The DNA pellet was washed in ice cold 70% ethanol and recentrifuged at 10000 r.p.m. for 10 minutes. After removing the ethanol, the pellet was dried under vacuum in a desiccator and dissolved in TE to the desired concentration.

2.2.2A(iii) Trichloroacetic Acid (TCA) Precipitation:

The proportion of [α -³²P] dATP incorporated into DNA following nick translation was determined by TCA precipitation as follows.

The nick translation reaction was diluted with 50 μ l of TE and 1 μ l was spotted onto the centre of a GF/C glass-fibre disc in 0.1ml of a salmon sperm DNA solution (0.5mg/ml in 20mM EDTA). 5ml of ice-cold TCA was then added and kept on ice for 10 minutes. The resultant precipitate was collected by filtering the solution through a second GF/C glass-fibre disc and washed 5 times with 5ml of ice-cold 10% TCA followed by 5ml of 96% ethanol. Both filters were dried and placed in liquid scintillation vials. The incorporated [α -³²P] was counted in an Interchange SL-30 scintillation counter. The first filter measured the total radioactivity in the sample while the second measured that incorporated into the DNA. The specific activity of the labelled DNA was determined by the c.p.m./ μ g of the DNA.

2.2.2A(iv) Scintillation Counting:

An Interchange SL-30 scintillation counter was used by employing the ¹⁴C channel set for ³²P counting. Scintillation fluid was prepared by mixing 660ml toluene AR grade, 330ml Triton X-100, 5g PPO and 0.2g POPOP. The filters were placed in plastic scintillation vials with 5ml of scintillation fluid. Counting efficiency was established by calculating disintegrations per minute (dpm) expected and comparing to the counts per minute obtained. In most cases the efficiency was found to be greater than 50%.

2.2.2A(v) Spectrophotometry:

The DNA concentration was calculated by measuring the absorbance (O.D.) at 260nm and 280nm in a Car 210 dual beam recording spectrophotometer, using a path length of 1cm. An O.D. of 1.0 at 260nm indicates a DNA concentration of 50µg/ml and an O.D. ratio of 1.8 for 260nm:280nm wavelengths indicates that the DNA preparation is pure and free from proteins and RNA.

The O.D. of bacterial cultures was measured at O.D. ⁵⁵⁰ in order to estimate the number of cells/ml. An O.D. of 1.0 at 550nm indicates a cell concentration of approximately 2×10^7 cells/ml for rec⁻ and 5×10^7 cells/ml for rec⁺.

2.2.2A(vi) Autoradiography:

Southern blots, M13 plaque lift blots, colony hybridisation blots, and DNA sequencing gels were autoradiographed on Fuji X-ray film in Harmer cassettes containing 2 intensifier screens at -70° C. Radioactively-labelled polyacrylamide gels were transferred to 3MM Whatman filter paper and autoradiographed wet at room temperature on Fuji X-ray film. Length of exposure was dependent on the intensity of the signal. Films were developed in a Kodak industrial X-OMAT processor-model 3.

2.2.2B DNA Preparation:

2.2.2B(i) Preparation of Covalently Closed Supercoiled Plasmid DNA:

This method was modified from Garfinkel *et al.* (1981). Cells were grown to an O.D. of 1.0 by shaking at 37° C in 500ml of 2X YT. Plasmids bearing the Col E1 origin of replication were amplified by adding $160\mu\text{g/ml}$ of chloramphenicol in the growing culture at an O.D. of 0.8 units. This arrests the chromosomal DNA replication but Col E1 derived plasmids continue to replicate and thus accumulate to a high copy number in the cell. The cells were then centrifuged in 500ml Sorvall bottles at 6000 r.p.m. for 10-15 minutes at 4° C. The cell pellet was resuspended in 4ml of TSE (25% sucrose, 50mM Tris.HCl, pH

8.0 and 50mM EDTA) which was followed by the lysis of the cells by adding 0.5ml of lysozyme solution (10mg/ml). This was incubated on ice for 10 minutes before the cells were completely lysed by the addition of 5ml of lytic mixture (2% Triton X-100, 50mM Tris.HCl, pH 8.0 and 50mM EDTA). Successful lysis is indicated by clearing of the solution which also becomes very mucillaginous. After incubation for 15 minutes on ice, the lysate was cleared by centrifugation for 90 minutes at 20000 r.p.m. in a Beckman Type 30 rotor at 14 °C. The bacterial debris and the chromosomal DNA were pelleted leaving mostly plasmid DNA in the supernatant.

The supernatant was transferred gently to a 25ml test tube containing 0.5ml of ethidium bromide (10mg/ml) and caesium chloride was added until the refractive index was 1.3957 (corresponding to a density of 1.66g/ml). This mixture was centrifuged for 15 minutes at 15000 r.p.m. in 40ml Sorvall tubes. The lipid layer, which rises to the top, was easily removed. The preparation was centrifuged in a VTi 65.2 rotor using quick seal tubes at 48000 r.p.m. and at 15 °C for 14-18 hours. This generated an equilibrium gradient in which different DNA species band at different heights corresponding to their buoyant density. These can be visualised by short exposure to long wave u.v. light.

The denser, lower band (covalently closed circular plasmid DNA) was collected by using a 19G needle connected to a

2ml syringe. The top of the tube was punctured first before the needle was gently introduced into the lower band by raising the tube on a jack. It was usually desirable to "reband" the DNA in caesium chloride/ethidium bromide by repeating the centrifugation in order to achieve greater purity.

Ethidium bromide bound to the DNA was removed by extracting with n-butanol (saturated with 20X SSC). The butanol sequesters the ethidium bromide, which can be removed as the top layer. This step was repeated 3-4 times. The DNA was then dialysed against 4 litres of TE with three changes over 24 hours at 4 C. The concentration of the DNA was measured as described in section 2.2.2A(v).

2.2.2B(ii) Quick Plasmid DNA Preparation:

This method is a modification of that described by Birnboim and Doly (1979). The four solutions required in this procedure are described below;

Solution I : 10mM EDTA, 25mM Tris.HCl, pH 8.0

Solution II : 0.2M NaCl, 1% SDS (always prepared fresh)

Solution III : 3M NaOAc, pH 4.8

Solution IV : 100mM NaOAc, 50mM Tris.HCl, pH 8.0

The cells from a 2ml overnight culture were pelleted in a microfuge for 1 minute and the supernatant removed by aspiration. The pellet was resuspended in 0.1ml of

solution I and kept on ice for 5 minutes. 0.2ml of solution II was gently mixed with the cell suspension and the cells left for a further 5 minutes at 0 C. After lysis, 1.5ml of solution III was added and left on ice for 10 minutes. Cell debris and chromosomal DNA were pelleted by centrifugation at 10000 r.p.m. for 5 minutes at 4 C. The supernatant was carefully transferred to a clean tube. Two volumes of absolute ethanol were added and the DNA precipitated (at -70 C for 10 minutes) and pelleted by centrifugation at 10000 r.p.m. for 10 minutes. The pellet was resuspended in 0.1ml of solution IV, and 2 volumes of absolute ethanol added for precipitating the DNA as described previously. The solution was spun at 10000 r.p.m. for 10 minutes in the microfuge. The pellet was washed with 70% ethanol, kept at -70 C for 5 minutes and centrifuged again at 10000 r.p.m. for 10 minutes. This washing was repeated and the pellet was dried under vacuum in a desiccator. The DNA was resuspended in 0.1ml of TE. 10µl of DNA solution were commonly used for analysis on an agarose gel or for restriction endonuclease analysis.

2.2.2B(iii) Preparation of Chromosomal DNA from E.coli and Other Bacterial Species:

This method was adapted from Garfinkel et al. (1981) and produces high molecular weight DNA suitable for digestion with restriction endonucleases.

One bacterial colony was grown overnight with shaking at 37 °C in 5ml of 2X YT media. After harvesting, the cells were washed in 1ml of solution of 50mM Tris.HCl, pH 8.0 and 20mM EDTA. The cells were spun in a microfuge for 2 minutes and resuspended in 0.3ml of the same solution. 0.05ml lysozyme (10mg/ml) was added and the cells were incubated on ice for 10 minutes. This was followed by the addition of 0.1ml of 10% SDS solution and a final concentration of 50µg/ml of proteinase K (dissolved in 50mM Tris.HCl, pH 8.0, 20mM EDTA and 1% SDS) to the preparation. The mixture was incubated at 65 °C for 10 minutes before gently extracting with phenol. The preparation was then centrifuged for 5 minutes at 5000 r.p.m. to separate the layers. The upper aqueous layer was collected using a very wide bore pipette to minimise shearing of the chromosomal DNA. The DNA was extracted a further 3 times with phenol as described in section 2.2.2A(i). The DNA was then precipitated with 2 volumes of absolute ethanol and collected by spooling, washed in 70% ethanol and dried under vacuum. In order to remove any contaminating RNA, the DNA was re-dissolved in 0.5ml of TE and incubated for 30 minutes at 37 °C with 0.05ml of preboiled RNase A (10mg/ml). After a further phenol extraction, the DNA was precipitated with absolute ethanol as described previously and the DNA was dried under vacuum before being resuspended in 0.2ml of TE. The DNA solution was stored at -20 °C.

2.2.2B(iv) Isolation of M13 Single-stranded DNA:

The method used for the preparation of M13 single-stranded DNA was according to Schreier and Cortese (1979), with the following modifications.

E.coli strain TG2 was grown to an O.D. of 0.2 units in 2X YT. Using a pasteur pipette, an isolated colourless/M13 plaque was picked into 2ml of the growing culture. The culture was vigorously shaken at 37 °C for 6-7 hours. After incubation, the cells were spun down in an eppendorf tube at 10000 r.p.m. for 2 minutes. 1ml of the supernatant fraction was transferred to another eppendorf tube containing 0.2ml of 2.5M NaCl, 20% polyethylene glycol 6000 and left on ice for 10 minutes. After 5 minutes centrifugation in a microfuge at 10000 r.p.m., all the supernatant fraction was drained from the tube. Following a short centrifugation, the last few drops of PEG-6000 were removed by using a drawn out pasteur pipette (formed by melting the pipette over a Bunsen burner). The phage pellet was then thoroughly dissolved in 0.1ml of TE (10mM Tris.HCl, pH 8.0, 0.1mM EDTA, pH 8.0) and mixed with 0.1ml of phenol. The DNA extraction procedure using phenol was the same as described in section 2.2.2A(i). The single-stranded DNA was precipitated in 0.3M sodium acetate, pH 6.0, with 2.5 volumes of absolute ethanol; the rest of the procedure, including the washing step, was as described previously in section 2.2.2A(ii). The DNA pellet was resuspended in 25µl of distilled water and stored at -20 °C until needed.

2.2.2B(v) Isolation of M13 Replicative Form DNA:

The following method was used for the preparation of Replicative form (RF) DNA of M13, derived cloning vectors and M13 derived recombinant phages.

A single purified plaque of M13 phage was picked into 1.5ml of a TG2 culture grown at 37 °C in 2X YT to an O.D. of 0.2 units. After 4 hours of growth at 37 °C, the culture containing M13 infected cells was centrifuged and the supernatant used to infect 500ml of growing TG2 at an O.D. of 0.2 units. These cells were then grown at 37 °C for 4 hours and then spun down. The M13 RF DNA was prepared from the cell pellet by the same method described for the preparation of the covalently closed supercoiled plasmid DNA [section 2.2.2B(i)].

2.2.2B(vi) Purification of DNA Restriction Fragments from Agarose Gels:

A special electro-elution box (figure 2.1) for eluting the DNA from agarose gels was made by the Imperial College Biochemistry workshop.

An agarose gel slice containing the DNA was cut into small pieces (2mm x 2mm) and placed in the sample wells in the centre of the electro-elution tank. The tank was filled with 500ml of 1X electro-elution buffer (10mM Tris.HCl, pH 8.0, 5mM NaCl, 1mM EDTA). All air bubbles were carefully

removed using a 0.1ml micropipette connected to a 1ml disposable syringe. 60µl of dye mix (3M NaOAC, pH 7.9, 0.025% bromophenol blue) were carefully added to the extraction holes. The DNA was eluted at 150 V for 30 minutes and at the end of the run, 0.4ml of the sample was withdrawn from the extraction hole via a 0.1ml micropipette and 1ml syringe. This procedure was repeated.

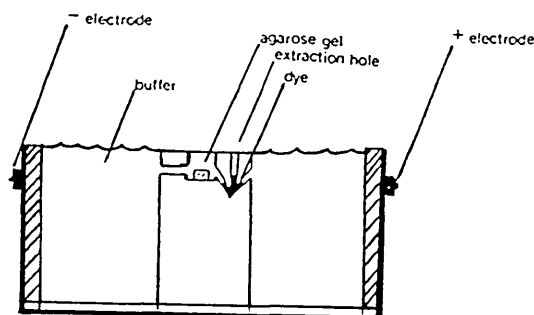


Figure 2.1 Electro-elution tank.

Two volumes of absolute ethanol and 0.1µg of glycogen were added to the sample, and placed in a dry-ice/isopropanol bath for 10 minutes. Recovery of the DNA was achieved by centrifugation in a microfuge at 10000 r.p.m. for 15 minutes. The DNA pellet was washed with 70% ethanol as described previously and then dried under vacuum in a desiccator. The DNA was resuspended in a small amount of distilled water and the concentration of the DNA determined by running an aliquot on a 0.8% agarose gel with known concentrations of DNA and visually comparing the fluorescent intensities of the DNA bands.

2.2.2B(vii) Purification of DNA Restriction Fragments from Acrylamide Gels:

The method described here was used for the preparation of small DNA fragments to be used as radioactive probes or for cloning into other vectors.

10µg of DNA were digested with appropriate restriction endonucleases (end filled with radioactive nucleotide if necessary) and loaded onto a 6% acrylamide gel. After electrophoresis the gel was visualised by u.v. (if stained with ethidium bromide) or autoradiographed (if radioactive) and the desired fragment cut out of the gel. The DNA was eluted overnight by soaking the gel in 0.5ml of distilled water. Cerenkov counting was performed at all stages to monitor the binding and elution of the radioactive DNA. The DNA was then extracted with phenol, precipitated with ethanol, washed in 70% ethanol and, finally, resuspended in 10µl of TE.

2.2.2C Gel Electrophoresis:

All gel electrophoresis was carried out in 1X TBE buffer. Gel loading dye (6X) was prepared with 0.25% bromophenol blue, 0.25% xylene cyanol and 40% (w/v) sucrose in distilled water.

Denaturing gel loading dye used was 90% formamide, 0.02% bromophenol blue and 0.02% xylene cyanol in 1X TBE buffer.

2.2.2C(i) Useful Molecular Markers:

Bacteriophage lambda DNA, digested with HindIII or with BstE II, and ϕ X 174 DNA, cleaved with HaeIII, were commonly used molecular markers to estimate the molecular weights of unknown DNA restriction fragments. The fragment sizes shown below are given in base pairs.

Lambda HindIII: 23130, 9416, 6557, 4361, 2322, 2027, 564, 125.

Lambda BstEII: 8454, 7242, 6396, 5686, 4822, 4324, 3647, 2323, 1929, 1374, 1264, 702, 224.

ϕ X174 HaeIII: 1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118, 72.

2.2.2C(ii) DNA Electrophoresis on Horizontal Agarose Gels:

Electrophoresis was carried out in BRL gel boxes (35 x 11.5 cm) at constant voltage between 20 and 100 V. The agarose gel was poured by melting agarose in 1X TBE at 95 °C for 5 minutes, followed by the addition of ethidium bromide (10mg/ml) to a final concentration of 0.5µg/ml. When the gel had set (this takes about 20 minutes) the comb was removed and the gel submerged in 1X TBE. After electrophoresis the gels were viewed on a 260nm transilluminator and photographed using a Polaroid MP4 land camera through an orange Kodak wratten filter.

2.2.2C(iii) Polyacrylamide Gels:

An acrylamide stock solution of 38% acrylamide and 2% bisacrylamide in distilled water was prepared. Once dissolved, the stock was de-ionised by mixing the solution with 20g of amberlite mixed-bed resin for 2 hours (later removed by filtration). A fresh solution of 10% ammonium persulphate was used for polymerisation.

All gels were run vertically between 20 x 40cm glass plates in 1X TBE at constant current.

2.2.2C(iv) Non-denaturing Thin Acrylamide Gels:

The 0.1mm thick gels were used for isolating small DNA fragments whose sticky ends have been filled with (³²P)dNTP.

5% polyacrylamide gels were used and prepared by mixing 0.4ml of 10X TBE, 0.5ml acrylamide:bisacrylamide 38:2 solution, 29.5ml distilled water, 400µl 10% ammonium persulphate solution and 70µl of TEMED. The gels were run at a constant current of 15 mA.

The distance migrated by a DNA fragment of a given size is dependent on the percentage of acrylamide present in the gel.

2.2.2C(v) Denaturing Gradient Acrylamide Gels:

These gels were used first to determine the melting behaviour of the target DNA attached to a GC rich region (GC clamp), and then to find the mutated fragment from the gel which have altered mobilities.

An ATTA, vertical electrophoresis system, with 13.5 x 13.5mm glass plates and 0.6 mm spacers, was used for these gels.

Two stock solutions were prepared;

0% Stock: 6.5% acrylamide
1X TBE buffer

80% Stock: 6.5% acrylamide
1X TBE
5.6M urea
32% formamide

In order to determine the melting behaviour of an unknown fragment, a perpendicular gradient of 0% to 80% denaturant was used. 80% denaturant solution was poured initially into the chamber of the gradient maker to ensure the exit of 80% followed by the 0% denaturant solution which was poured into the empty chamber. The liquid gels were allowed to pass by gravitational flow through thin tubing into the glass plates, placed perpendicular to the direction of electrophoresis. After polymerisation was complete, the DNA samples were uniformly loaded across the top well of the

gel in a volume of 100-200 μ l non-denaturing loading buffer (25% glycerol, 1mM EDTA and 0.5% bromophenol blue: xylene^ocyanol). The gel was run in a 60 C oven at 150 V, in preheated 1X TBE buffer with a peristaltic pump connected for better circulation of the buffer.

Electrophoresis was continued for 4-6 hours before the gel was removed and stained with 1 μ g/ml ethidium bromide (5 mg/ml) to be examined under uv illumination. The GC clamp did not melt at the temperature and denaturant concentrations used in these experiments.

After the gel conditions were optimised for the target fragment attached to the GC clamp, the DNA preparations from the pooled mutagenesis step [see section 2.2.3B(ii)], were subjected to preparative denaturing gradient gel electrophoresis for purification of the mutant DNA fragments from the wild-type.

The preparative gel was poured parallel to the direction of electrophoresis and in the same manner as described previously except that the acrylamide solution was dripped into the plates from the top rather than from the sides.

4-6 μ g of DNA sample (digested with the appropriate enzymes) were loaded into a 6mm wide lane and the gel electrophoresed as described previously. At the end of the run, the gel was stained with ethidium bromide as before. In the case of nitrous acid treated samples, a very faint

band above the wild-type band was just visible. However, these regions were carefully excised with a razor blade from the gel and the DNA eluted as described in section 2.2.2B(vii).

2.2.2C(vi) Non-denaturing Gradient Thick Acrylamide Gels:

These gels were used for isolating small sized DNA fragments of less than 100 bp. The gels were prepared as for denaturing gels except that no urea was added to the polymerisation mixture and 2.5mm spacers were used. A stock solution of 38% acrylamide:2% bisacrylamide was used. The gels were electrophoresed in 1X TBE and at 15-20 mA and stained with 1mg/ml ethidium bromide (5mg/ml).

2.2.2C(vii) Denaturing Thin Acrylamide Gels for DNA Sequencing:

The gel plates were first cleaned with ethanol and then treated with dimethyl-dichloro-silane solution. 0.35mm thick strips of "plasticard" (Raven Scientific Ltd.) were placed along the longitudinal edge of one of the plates and the two then taped together. The gel was prepared with 25g of "ultrapure" urea, 350µl of 10% ammonium persulphate, 5ml of 10X TBE solution, 7.5ml of 40% (w/v) of acrylamide stock and distilled water to a volume of 50ml, giving a final concentration of 6% acrylamide. The solution was filtered through a Nalgene 20µm filter unit and degassed

prior to use. Polymerisation was initiated by the addition of 70µl of TEMED. The gel was poured between two plates, and a plasticard comb inserted before polymerisation was complete.

2.2.2C(viii) Analysis of Proteins on 12.5% Discontinuous SDS Polyacrylamide Gels:

The method for running these gels was of Laemmli (1970) and the gels were prepared as follows:

Acrylamide/bisacrylamide solution: 30g of acrylamide and 0.8g of bisacrylamide were dissolved in 100ml of distilled water and filtered through a Whatman No.1 filter.

Resolving gel buffer: 3M Tris.HCl, pH 8.8.

Stacking gel buffer: 0.5M Tris.HCl, pH 6.8

Reservoir buffer: 0.025M Tris, 0.192M glycine, pH 8.3.

Resolving gel: 12.5ml acrylamide/bisacrylamide solution, 3.75ml resolve stock, 0.3ml 10% SDS, 1.5ml 1.5% ammonium persulphate (APS) and 11.95ml distilled water were degassed for 1-2 minutes and 20µl of TEMED were added before pouring the gel.

Stacking gel: 2.5ml acrylamide/bisacrylamide solution, 5ml stack stock, 0.2ml 10% SDS, 1ml 1.5% APS and 11.3ml distilled water were degassed for 1-2 minutes and 20µl of TEMED added before pouring the gel.

The gel was run at 120 V to stack and 200 V to resolve, or 50 V overnight, in reservoir buffer until the bromophenol blue dye reached the bottom of the gel.

The gel was stained in 0.2% Coomassie blue (Brilliant blue dissolved in 95% ethanol) and fixed in 20% acetic acid for 2 hours. Destaining was carried out in 20% ethanol/3.5% acetic acid for 4-5 hours with Amberlite monobed resin and synthetic sponge pieces (to absorb the excess Coomassie blue dye).

2.2.2D Modification of DNA by Enzyme Catalysis:

2.2.2D(i) Restriction Endonuclease Digestion of DNA:

Most restriction endonucleases were used in the presence of one of four buffers (1X) and incubated at 37 °C for 1-2 hours. 10X reaction buffers were prepared and stored in small aliquots at -20 °C until required.

10X low salt buffer: 100mM Tris.HCl, pH 7.4; 100mM MgSO₄; 10mM DTT.

10X medium salt buffer: 500mM NaCl; 100mM Tris.HCl, pH 7.4; 100mM MgSO₄; 10mM DTT.

10X high salt buffer: 1M NaCl; 500mM Tris.HCl, pH 7.4; 100mM MgSO₄; 10mM DTT.

10X SmaI buffer: 200mM KCl; 100mM Tris.HCl, pH 8.0; 100mM MgCl₂; 10mM DTT.

Digestions were terminated by a phenol extraction step followed by ethanol precipitation of the DNA as described in sections 2.2.2A(i) and 2.2.2A(ii).

2.2.2D(ii) In vitro Ligation of DNA Using T4 DNA Ligase:

Ligation of DNA fragments were performed according to Maniatis et al. (1982).

Cohesive end ligations were carried out in 66mM Tris.HCl, pH 7.5; 1mM ATP; 1mM spermidine; 10mM MgCl₂; 15mM DTT and 0.2mg/ml BSA for 12-16 hours at 15 °C with 5 units of T4 DNA ligase per µg of DNA.

Blunt end ligations were performed under the same conditions but with 50 units of enzyme per µg of DNA and in the presence of 15% PEG 6000.

The enzymes were inactivated by a phenol extraction step before precipitation of the DNA. 10-20ng of ligated DNA were routinely analysed by electrophoresis on a 0.8 % agarose gel to check the efficiency of ligation.

2.2.2D(iii) Incorporation of Radioactive Nucleotide into DNA by Nick Translation:

Nick translation reactions were carried out as described by Rigby et al. (1977) using a kit purchased from Amersham International. 50µCi of [alpha-³²P] dATP were used per

200ng DNA at the reaction conditions recommended by Amersham. DNA was routinely labelled to a specific activity of 5.0 to 7.5E 8 dpm/ μ g and used directly as a hybridisation probe after boiling.

2.2.2D(iv) Incorporation of [γ -³²P] ATP at the 5' End of Oligonucleotides Using T4 Polynucleotide Kinase:

Synthetic oligonucleotides containing a free 5' hydroxyl group were phosphorylated with [γ -³²P] ATP directly. The labelling reaction included 10 picomoles of the oligonucleotide, 2 units of T4 polynucleotide kinase in 70mM Tris.HCl, pH 7.5; 10mM MgCl₂; 5mM DTT in a total volume of 100 μ l at 37 °C for 1 hour. After incubation, the labelled DNA was separated from unincorporated radionucleotide by centrifugation at 1000 r.p.m. for 4 minutes through a Sephadex G-50 (Pharmacia) column (0.9ml bed volume) equilibrated with 10mM Tris.HCl, pH 8.0; 100mM NaCl; 1mM EDTA (Maniatis et al., 1982).

2.2.2D(v) Dephosphorylation of the 5' Terminus of DNA:

This was accomplished using calf intestinal phosphatase (CIP). After digestion with the appropriate restriction enzymes, 1 μ g of DNA was incubated with 2 units of CIP in 20 μ l reaction volume containing 1X CIP buffer (50mM Tris.HCl, pH 9.0; 1mM MgCl₂; 0.1mM ZnCl₂ and 1mM spermidine)

at 37 °C for 15 minutes, then at 56 °C for a further 15 minutes. The sample was then treated with 10mM nitriloacetic acid for 1 hour at 65 °C, phenol extracted, ethanol precipitated, washed with 70% ethanol and resuspended in 10µl of TE. To ensure that the dephosphorylation was successful, the ligation reaction using the vector alone and T4 DNA ligase was used as a control (no ligation should occur).

2.2.2D(vi) Preparation of Labelled DNA Restriction Fragments Using DNA Polymerase 1, Klenow-Fragment:

The 5' overhanging end of the restriction endonuclease digested DNA was filled in using [α -³²P] dCTP and Klenow fragment.

100ng of DNA were incubated in a 20µl reaction volume containing 50mM Tris.HCl, pH 7.2; 10mM MgSO₄; 0.1mM DTT; 50µg/ml BSA; 0.2mM of all dNTPs; 2µci of [α -³²P] dCTP and 1 unit of Klenow enzyme, at room temperature for 30 minutes. The reaction was stopped by the addition of EDTA to a final concentration of 5mM.

2.2.2D(vii) Unidirectional Deletion Analysis Using Exonuclease III and S1 Nuclease:

5µg of R F DNA of M13 phage vectors VHV12 and VHV22 were cleaved with BamHI (3' recessed end) and PstI (3'

overhanging end) REs until completion. The DNA was extracted with phenol, phenol-chloroform mixture (1:1), ethanol precipitated, washed with 70% ethanol and dissolved in 100µl of distilled water. 10µl of 10X exonuclease buffer (500mM tris.HCl, pH8.0, 50mM MgCl₂ and 100mM DTT) and 100 units of exonuclease III were added to the DNA solution. This solution was incubated at 30 °C and 10µl aliquots were then taken at 1 minute and 10 minute intervals in the case of VHV12 and VHV22 respectively.

Samples were transferred to eppendorf tubes containing 100µl of 1X S1 nuclease buffer (30mM NaOAc, pH 4.6, 50mM NaCl and 1mM ZnSO₄) and kept on ice until the last aliquot was collected. 1 unit of S1 nuclease was added to each tube and incubated on ice for 30 minutes. The DNA was phenol extracted, ethanol precipitated, resuspended in 10µl of distilled water and 1µl of each sample loaded onto a 0.8% agarose gel. 1µl of the DNA was ligated to recircularise and transformed to bacterial strain TG2. M13 plaques were screened by DNA sequencing reactions (ddTTP track).

2.2.2D(viii) DNA Sequencing:

The method employed was that of Sanger et al. (1977) with the following modifications.

DNA fragments were first cloned in M13 vectors mp8 and mp10 (Messing and Vieira, 1982) which contain multiple cloning sites in the alpha-peptide of the beta-galactosidase gene

(see section 2.2.2E). Single-stranded DNA was prepared from mature phage particles and used directly in the sequencing reaction as follows.

The sequencing reaction consisted of a working solution, chase solution, M13 universal primer, ³⁵S-dATP (2000 Curie/mM) and DNA polymerase 1, large fragment (Klenow).

2.2.2D(viia) DNA Sequencing Reaction Mixture:

The working reaction mixture for use with ³⁵S-dATP was made as follows;

	T	C	G	A
0.5mM dTTP (μl)	6.25	125.00	125.00	125.00
0.5mM dCTP (μl)	125.00	6.25	125.00	125.00
0.5mM dGTP (μl)	125.00	125.00	6.25	125.00
10.0mM ddTTP (μl)	12.50	-	-	-
1.0mM ddCTP (μl)	-	18.00	-	-
10.0mM ddGTP (μl)	-	-	5.00	-
1.0mM ddATP (μl)	-	-	-	5.00
dds water (μl)	230.00	225.00	240.00	120.00

Chase solution used was 0.25mM dNTP in distilled water.

2.2.2D(viib) Annealing and Polymerisation:

5μl of DNA template, 1μl (2.5pmol) M13 universal primer, 1.5μl of 10X sequencing buffer (10mM MgCl₂, 10mM NaCl) and

4µl of distilled water were mixed in a 0.5ml eppendorf tube and covered with a thick layer of aluminium foil. This template-primer complex was placed at 65 °C in a water bath for 20 minutes, then removed along with the water of the water bath and allowed to cool slowly to room temperature. 1.5µl of ³⁵S-dATP and 2 units of Klenow were added to the annealed template and spun down in a microfuge for a few seconds. After mixing, 2.5µl aliquots of this mixture were placed in 4 eppendorf tubes labelled A, G, C, and T containing the appropriate working sequencing solution. The reaction was allowed to proceed for 20 minutes at 37 °C after which time 2µl of the chase solution were added to each tube and incubation continued for a further 15 minutes at 37 °C. 4µl of the stop solution were added to each tube.

2.2.2D(viic) Gel Loading:

An aluminium plate (3mm thick and of the same dimensions as the gel plate) was always placed on the back of the glass plate and the gel pre-run at 40 Watts (in order to obtain even heating) for 45-60 minutes before loading. 6% denaturing acrylamide gels were used routinely and the order of loading was always A, G, C, T. Before loading the samples were heated to 95 °C for 2 minutes. Whilst the samples were still hot, 2µl of sample per track were loaded via a fine drawn-out pasteur pipette. In some cases, the

samples were loaded at the same intervals on different gels and electrophoresed for 2 and 5 hours separately. After electrophoresis the gel was transferred to 3MM Whatman paper and dried on an ATTO Gel drying processor (AE-3700) at 80 C^o for 40-60 minutes and then set for autoradiography as described in section 2.2.2A(vi).

2.2.2E Cloning into M13 Phage Vectors:

M13 is a filamentous male-specific phage with a single-stranded genome that replicates via a double-stranded intermediate. Messing and Viera (1982) have constructed a series of phage vectors by cloning the alpha peptide of beta-galactosidase into the replicative form DNA of M13 phage. A polylinker site was later introduced into the coding region of beta-galactosidase to facilitate cloning. Wild-type and recombinant clones can be selected on a bacterial strain having a deletion in the alpha peptide of beta-galactosidase. When X-gal, a chromogenic substrate for beta-galactosidase, is added to the medium in the presence of inducer IPTG any wild-type phage will produce blue plaques upon the cleavage of X-gal by the complementation of the alpha and beta peptides of the beta-galactosidase; recombinant phages will have a white morphology. However, a correct transcriptional read-through into the cloned fragment may produce blue plaques with a very low frequency.

2.2.2E(i) Identification of the insert in M13 phage vectors:

After ligation of the restriction fragments into the appropriate sites of the M13 phage vectors, 100ng of the ligated DNA were used to transform into bacterial strain TG2. Transformants were then added to 3ml of molten 2X YT top agar placed at 45 °C, together with 200µl of an exponentially growing culture of TG2 and 50µl of IPTG (20mg/ml in distilled water) and 50µl of X-gal (25mg/ml in dimethylformamide). This mixture was vortexed and quickly poured onto 2X YT plates, allowed to set and incubated at 37 °C. White plaques were picked for further analysis.

2.2.2E(ii) Identification of M13 Clones with Insert in Opposite Orientation:

This method was used to identify M13 clones containing insertions in the opposite orientation to each other. 50µl of M13 phage supernatant of test and reference clones were mixed in an eppendorf tube containing 20µl of 20% PEG 6000 in 2.5M NaCl, and kept on ice for 5 minutes. In separate tubes, test and reference clones were added separately to the same solution. The phage particles were spun down in a microfuge at 10000 r.p.m. and the pellet was resuspended in 20µl of distilled water, 3µl of 10% SDS plus 6µl of 20X SSC. The mixtures were incubated at 67 °C for 10 minutes to allow melting and then cooled slowly to room temperature to

rehybridise. The samples were loaded onto a 1% agarose gel for electrophoresis [see section 2.2.2C(ii)]. Samples containing inserts, as small as 40bp, in opposite orientation show reduced mobility as a result of DNA hybridisation.

2.2.2F In vitro Hybridisation Techniques:

For in vitro hybridisation studies the following stock solutions were prepared;

- (1) Formamide - 100% deionised w/10g Amberlite. Stored in 200ml aliquots at -20^o C.
- (2) 20X SSC.
- (3) 1M Sodium phosphate buffer, pH 6.5 .
- (4) Salmon sperm DNA - 5mg/ml.
- (5) 50X Denhardt's solution - 1% Ficoll, 1% Polyvinyl pyrrolidone, 1% BSA in distilled water.
- (6) 50% Dextran sulphate solution.

Denaturing solution: 1.0M Tris.HCl, pH 8.0; 0.5M NaOH.

Neutralising solution: 1.5M NaCl; 1.0M Tris.HCl, pH 8.0.

Prehybridisation Mix (10ml): 0.1g glycine; 1ml salmon sperm DNA (denatured); 0.5ml 1M sodium phosphate buffer; 1ml 50X Denhardt's solution; 2.5ml 20X SSC; 5ml formamide.

Hybridisation Mix (10ml): 2.5ml 20X SSC; 2ml 50X Dextran sulphate; 5ml formamide; 200µl 1M sodium phosphate buffer;

0.2ml salmon sperm DNA (denatured); 200 μ l 50X Denhardt's solution; denatured probe.

2.2.2F(i) Southern Blot Analysis:

The localisation of particular DNA fragments in the total genomic or plasmid DNA was carried out by the transfer technique described by Southern (1975).

After photographing the agarose gel, 250ml (for 110 x 140mm gels) of denaturing solution were added and the gel was shaken gently for 30 minutes. The liquid was poured off and the gel rinsed in distilled water. 250ml of neutralising solution were added and the gel shaken for a further 30 minutes. If high molecular weight DNA was to be blotted, an additional step of depurination (soaking the gel in 0.25M HCl for 15 minutes) was employed before denaturing the gel. The DNA was transferred to a nitrocellulose filter by the blotting technique of Southern (1975). When blotting was completed, the origin was marked on the filter with black biro and then baked for 1 hour at 80 °C or placed on a uv transilluminator to uv cross-link for 5 minutes. 10ml of prehybridisation solution were added to the filter which was shaken gently at 42 °C for at least 2 hours. Nick-translated probe (boiled for 3 minutes) was added to the solution. Hybridisations were left overnight at 42 °C before washing. All washings were carried out at 65 °C for 2-4 hours either in stringent conditions (0.25X SSC, 0.5% SDS),

or in non-stringent conditions (1X SSC, 0.1% SDS). The filter was set up for autoradiography, whilst still slightly damp, for a suitable time according to the intensity of the bound radioactivity as described in section 2.2.2A(vi).

2.2.2F(ii) M13 Plaque Lift Hybridisation:

For 9cm petri dishes: 30ml LCM bottom agar plus 5ml LCM top agar containing 30 μ l X-gal (20mg/ml), 30 μ l IPTG (24mg/ml), and 100 μ l plating cells were used. For large petri dishes, three times these quantities were used.

Nitrocellulose filters cut to the size of the petri dish were used to transfer M13 phage from the top agar. The filters were placed in denaturing solution for 2 minutes to liberate and denature the DNA and subsequently placed in neutralising solution for 2 minutes. After air drying for 15 minutes on 3MM Whatman paper, they were baked at 80 °C for 1 hour. After prehybridising for 2 hours (in prehybridisation solution) at 42 °C, a labelled probe (100ng to four small filters or two large filters), previously denatured by boiling for 3 minutes, was added to the solution. The filters were incubated at 42 °C with gentle shaking overnight followed by washing in 2X SSC and 0.1% SDS at 65 °C for 2-4 hours.

2.2.2F(iii) In Situ Colony Hybridisation:

The method employed was essentially that of Grunstein and Hogness (1975). Bacterial colonies between 0.5 and 1.0mm in diameter were transferred directly from the petri dish to Schleicher and Schewell BA85 nitrocellulose membranes. The filters were placed colony-side-up on 3MM Whatman paper saturated with 10% (w/v) SDS for 3 minutes, then onto paper saturated with 1.5M NaCl plus 500mM NaOH, for 5 minutes. The filters were baked at 80 °C for 90 minutes. Prehybridisation, hybridisation and washing conditions were the same as those described for M13 plaque hybridisation [section 2.2.2F(ii)]. The filters were autoradiographed as described in section 2.2.2A(vi).

2.2.3 Base Substitution. Mutagenesis Techniques:

This section describes the strategies adopted for site directed (section 2.2.3A) and region directed random (section 2.2.3B) mutagenesis on the attTn7 DNA.

2.2.3A Site-specific, Mismatched Oligonucleotide Directed Mutagenesis:

A 5'TAATAAATCGATGCCCT 3' heptadecamer was purchased from Prof. A.R. Fersht's laboratory at Imperial College, London. The mismatch was deliberately placed in the centre in selection of the oligomer which is discussed in section 2.2.3A(iv). Centrally positioned mismatches give the

maximum binding differential between perfectly matched duplexes and mismatched duplexes (Zoller and Smith, 1983).

2.2.3((i) Purification of the Heptadecamer:

The crude oligonucleotide was dissolved in an equal volume of fresh ammonia solution (.880, AR), incubated overnight at 55 °C and dialysed against water to remove low molecular weight material which can interfere with the migration of the oligonucleotide in the gel. The dialysate was then lyophilised and dissolved in 0.5ml of distilled water. 20µl of the crude material were mixed with 20µl of 10mM EDTA in 98% deionised formamide and boiled for 2 minutes prior to loading onto a polyacrylamide gel as described below.

2.2.3A(ii) Preparation of Gel and Electrophoresis:

The heptadecamer was purified from low molecular weight oligonucleotide material using a 20% polyacrylamide gel. The sample was loaded onto a gel (20 x 40cm, 0.4mm thick) containing 10ml 5X TBE, 25ml 40% acrylamide stock and 21g urea in a total volume of 50ml. After electrophoresis at 20 mA for 90 minutes, one gel plate was carefully removed and Saran Wrap placed on top of the gel. The second plate was then removed, ensuring that the gel was still attached to the Saran Wrap. A fluorescent thin layer chromatography plate was placed beneath the gel and illuminated from

above by uv. shadowing at 254nm. A series of bands could be seen which corresponded to oligomers of decreasing size as the dye moved down the gel. The top band was carefully cut into small (2 x 2mm) pieces which were soaked overnight in 3ml of distilled water at 37 °C. After filtering the eluted oligonucleotide through siliconised glass wool to remove gel debris, the filtrate was dialysed overnight against 1 litre of distilled water with three changes.

2.2.3A(iii) Estimation of Oligonucleotide Concentration:

The concentration of the oligonucleotide was determined spectrophotometrically according to Zoller and Smith (1983).

The relationship between ϵ_0 , OD_{260} and concentration of oligonucleotides is expressed as $OD_{260} / \epsilon_0 = xM$ where ϵ_0 is the millimolar extinction coefficient of the oligonucleotide. This is determined by summing the contribution of each base (where $\epsilon_0 T = 9.7$, $C = 9.2$, $G = 11.7$, $A = 15.4$ per mole).

For the oligo (described in section 2.3.3A), the ϵ_0 was determined to be 200.5, the OD_{260} to be 2.27, and therefore, the concentration was calculated to be 1.13E-4mM or 11.3 picomol per μ l.

2.2.3A(iv) Use of Oligonucleotide as a Mutagenic Primer:

200ng of single-stranded DNA (MAS to be mutagenised) was annealed to 2.5 pmole each of purified oligonucleotide and M13 universal primer by the addition of 10mM MgCl₂ and 5mM NaCl, heating to 65 °C for 15 minutes and cooling to room temperature. The annealed template was then synthesised using 5 units of DNA polymerase Klenow, in the presence of 200µm dNTP's (all four) and 1 unit of T₄ DNA ligase in 20µl volume. This was followed by the addition of 5 units of Klenow for another 1 hour. The DNA was extracted with phenol and precipitated with ethanol, as described in sections 2.2.2A(i) and 2.2.2A(ii) respectively, and then treated with 1 unit of S1 nuclease in S1 buffer as described in section 2.2.2D(vii).

10ng of DNA were transformed to competent 71/18 mutL strain as described in section 2.2.1(i) and the M13 plaques were screened for a mutant as follows.

The plaques were transferred to nitrocellulose filters and prepared for prehybridisation as described in section 2.2.2F(ii). The filters were treated with prehybridisation solution (0.9M NaCl; 0.09M Tris.HCl, pH 7.4; 6mM EDTA; 0.5% sodium phosphate buffer; 2X Denhardt's solution and 0.2% SDS) at 37 °C for 2 hours. The 5' [gamma-³²P] dATP labelled oligonucleotide was then added to the filters and incubated for 10 hours at 37 °C. The filters were washed with 5X SSC at room temperature for 1 hour and then autoradiographed.

2.2.3B Region Directed Random Mutagenesis Techniques:

Three methods were adopted for obtaining base substitutions in the 70 bp attTn7 site. The two methods which failed to yield mutants with high frequency will be discussed first, followed by the successful method.

2.3B.1 Forced Misincorporation:

Forced misincorporation of nucleotides during DNA strand repair in vitro was performed with a modification of the method of Zakour and Loeb (1982). 5µg of CCC DNA were digested with 10 units of BamHI RE at room temperature for 30 minutes, in the presence of 5µg ethidium bromide, in a 100µl reaction volume in order to generate a specific nick in each strand at the BamHI RE site. The DNA was checked on a 0.8% agarose gel [the open circular (OC) molecule had retarded mobility in comparison to CCC and linear molecules]. The DNA was extracted with phenol, phenol:chloroform (1:1), precipitated with ethanol and resuspended in 100µl of TE.

20µl of DNA solution was subjected to ExoIII treatment with 0.5 to 2.5 units/µg DNA at 30 °C for 1-5 minutes; samples were taken at one minute intervals and then extracted with phenol as described previously. The ExoIII treated DNAs repaired in vitro in four different tubes using 5 units of AMV reverse transcriptase in RT buffer (10mM Tris.HCl, pH 8.3; 1mM MgCl₂; 14mM KCl; 2mM beta

mercaptoethanol) in the presence of 200µM dNTPs; each tube lacked one of the dNTPs (i.e. dATP, dCTP, dGTP or dTTP). The reaction was carried out for 15 minutes at 37 °C and terminated by adding 5µl of 0.5M EDTA. DNA was extracted with phenol as described in sections 2.2.2A(i) and 2.2.2A(ii). Normal repair (i.e. containing 200µM of all 4 dNTPs) was also carried out by adding 5 units of fresh AMV reverse transcriptase for 15 minutes at 37 °C. After extracting the DNA with phenol, 10ng of DNA from each reaction were transformed into TG2. The transformants were screened by a DNA sequencing reaction (dATP, dGTP, dCTP and dTTP track analyses for reaction tubes without T, G, C and T respectively). Only two mutants out of 100 analysed were identified and these mutants were outside the region of interest.

2.2.3B(ii) Myers Mutagenesis Technique:

This section contains a description of a method for saturating cloned DNA fragments with single base substitutions (Myers et al., 1985). This technique failed to produce any point mutants, however, one deletion mutant was obtained.

In three separate reactions, approximately 40µg of single-stranded DNA were treated with (1) 60µl of 18M formic acid; (2) 60µl of 12M hydrazine and (3) 60µl of nitrous acid (10µl of 2.5M NaOAc, pH 4.3 and 50µl of 2M sodium nitrite)

in a final volume of 100 μ l. Reactions 1 and 2 were incubated for 20 minutes and reaction 3 for 120 minutes at room temperature. The reactions were then stopped by the addition of 200 μ l of 2.5M NaOAc, pH 5.5, 100 μ l of distilled water and 1ml of ethanol. The DNA was precipitated and washed as described in section 2.2.2A(i) and resuspended in 80 μ l of distilled water.

The DNA was mixed with 10 μ l of 10X KS buffer (70mM Tris.HCl, pH 7.5; 70mM MgCl₂; 0.5M NaCl and 20mM DTT) and 40 pmole M13 universal oligonucleotide primer. After heating to 75 $^{\circ}$ C for 5 minutes the DNA was annealed at 40 $^{\circ}$ C for 15 minutes. The chemically treated DNA had a second strand synthesised by adding 30 units of AMV reverse transcriptase and using all four dNTPs to a final concentration of 200 μ M.

After incubation at 40 $^{\circ}$ C for 1 hour, the DNA was extracted with phenol and precipitated with ethanol as described in sections 2.2.2A(i) and 2.2.2A(ii). The double-stranded target DNA fragment was purified by acrylamide gel electrophoresis as described in section 2.2.2B(vii). The purified DNA was ligated into an untreated plasmid pGC2 backbone. The ligation mixture was transformed into E. coli TG2_r and all Ap^r colonies were pooled to isolate plasmid DNA. 100 μ g of the pooled DNA were digested with appropriate endonucleases to excise the target fragment attached to the GC clamp. The DNA was precipitated with ethanol and resuspended in 20 μ l of non-denaturing loading buffer (25%

glycerol, 1mM EDTA and 0.5% bromophenol blue and xylene cyanol).

The DNA sample was loaded onto a denaturing perpendicular gradient acrylamide gel [see section 2.2.2C(v)]. The mutated DNA fragments (attached to the GC clamp) were eluted from the preparative denaturing gradient gel and ligated to the appropriate ends of the M13 phage vector, mp11. The DNA was transformed into E. coli TG2 and the transformants were screened by DNA sequencing using the dideoxy chain termination method (Sanger et al., 1977) as described in section 2.2.2D(viii).

2.2.3B(iii) Use of Sodium Bisulphite as a Mutagen to Introduce Base Substitution:

This method was used to generate base substitutions of C to T or G to A type in the attTn7 DNA and which was adapted from Shortle and Bostein (1983) with the following modifications.

2.2.3B.3(iiia) Preparation of the Gapped Circular Molecules for Target Mutagenesis:

For sodium bisulphite mutagenesis single-stranded DNA is required. In order to obtain the single-stranded target sequences in either the Watson or Crick strand of the mini attTn7 site, the following strategy was adopted.

10µg of RF DNA of M13 phage vectors MP11~~ATT~~GC and mp10 were digested to completion with HindIII and EcoRI REs respectively. The DNAs were extracted with phenol, washed with 70% ethanol and precipitated with absolute ethanol as described previously. 5µg of the digested MP11~~ATT~~GC and mp10 DNA's were mixed with 5µg of single-stranded MP11GC and ΔL849 (respectively) in sequencing buffer (10mM MgCl₂, 10mM NaCl) and a total volume of 50µl (see figure 8.5 for details). To obtain the gapped circular molecules, the DNA was denatured by boiling at 95 °C for 3 minutes and the tube containing DNA placed quickly on ice. This rapid annealing yielded three kinds of DNA molecules in the population; linear molecules with, or without, the target sequence and gapped circular molecules with a single-stranded target sequence.

2.2.3B(iiiB) Procedure for Sodium Bisulphite Treatment:

A solution of 4M sodium bisulphite, pH 6.0, was prepared freshly prior to use by dissolving 156mg of sodium bisulphite (NaHSO₃) and 64mg of sodium sulphite (Na₂SO₃) in 0.5ml of distilled water. Three volumes of this solution were added to one volume of DNA solution (50µg/ml in 15mM NaCl, 1.5mM sodium citrate, pH 7.0) in an eppendorf tube, followed by 0.04 volume of freshly prepared 50mM hydroquinone in a 0.5ml reaction volume. After thorough mixing, 0.2ml of paraffin oil was layered above the mixture

followed by incubation at 37 °C in the dark. 0.1ml aliquots of the incubation mixture were taken after 1, 2, 3, 4, and 10 hours of incubation.

The reaction was terminated, and sodium bisulphite removed, by dialysing the sample against 5000 volumes of 5mM potassium phosphate, pH 6.8 and 0.5mM hydroquinone with 2 changes at 4 °C. This was followed by dialysis against 5000 volumes of 5mM potassium phosphate, pH 6.8. To eliminate bisulphite from 5, 6 dihydrouracil 6-sulphonate at this stage, dialysis was continued against 2000 volumes of 0.2M Tris.HCl, pH 9.2, 50mM NaCl, 2mM EDTA at 37 °C for 24 hours. Dialysis was further continued in 2000 volumes of TE at 4 °C for 12 hours. The DNA was then recovered and concentrated by ethanol precipitation (see section 2.2.2A(ii)).

2.2.3B(iiic) Identification of the Mutants:

0.5µg of mutagenised DNA was repaired in vitro using 2 units of DNA polymerase Klenow-fragment, 200µM dNTP (i.e. dATP, dGTP, dCTP and dTTP) in a 20µl reaction volume. Before transforming the DNA into E. coli TG2, 1µg DNA was treated with 0.1 units of S1 nuclease for 5 minutes. A 10µl aliquot was transformed to competent TG2 [as described in section 2.2.1(i)] and the transformants were screened by DNA sequencing reactions (dGTP or dATP track only) for identification of the mutants. The disappearance of a single or multiple G band and the appearance of a single or

multiple A band on the sequence ladder confirmed base substitution(s) from G to A.

The complete DNA sequences of some mutant clones (picked at random) were determined in order to confirm that no other DNA rearrangements had been caused by the sodium bisulphite treatment.

2.2.4 In vitro DNA-Protein Interaction Studies:

The binding assays of the attTn7 DNA with the bacterial crude extracts were carried out according to Fried and Crothers (1983) with the following modifications.

2.2.4.(i) Preparation of Crude Extracts:

Bacterial strains TG2 (pUC18) and TG2 (pFK::Tn7) were grown in 100ml of LB media to an O.D. of 0.4 units and then centrifuged at 2000 r.p.m. for 10 minutes at 4 C. The cell pellets were resuspended in 10ml of solution C (50mM HEPES, pH 7.8; 50mM KCl; 0.1mM EDTA and 1mM DTT) and chilled on ice before passing through a "French Pressure Cell Press" at 12000 psi (pounds per square inch) with a selected ratio of 762. The fractions were collected and spun down immediately at 10000 r.p.m. for 10 minutes at 4 C. The supernatants were transferred to fresh tubes containing a final concentration of 1mM PMSF and 10% glycerol. 0.5ml aliquots of this fraction were snap freezed in liquid

nitrogen and then stored at -70°C . Protein concentrations in this fraction were determined by using the Bio-Rad assay kit and checked on a 12.5% discontinuous SDS-polyacrylamide gel as described in section 2.2.2C.(viii).

2.2.4(ii) Formation of DNA-Protein Complex In Vitro:

The 70 bp attTn7 DNA, the target for this study, was isolated from a gel by digesting the MASRF DNA with HindIII and EcoRI REs, and then end filled with radiolabelled dCTP using the Klenow fragment as described in section 2.2.2D(vi).

100ng of end filled DNA were mixed with 1, 5, 10, 20, 40, 50, 70 and 100 μl of crude extract separately and incubated at 4°C for 10 minutes. The reaction was stopped by adding 5mM EDTA and then immediately loading onto a 5% acrylamide gel without adding any dye. The gel was prepared as described in section 2.2.2C(iv) but with the addition of TE instead of TBE buffer. The electrophoresis was carried out at 20 mA overnight in 10mM Tris.HCl, pH 8.0 and 0.7mM EDTA using a peristaltic pump connected for buffer recirculation. After electrophoresis the gel was removed and autoradiographed for 10 hours.

CHAPTER 3

The Use of M13 Filamentous Phages as Transducing Particles for Determining the Frequency of Transposition of Tn7 in vivo

3.1 Introduction:

The aim of this project was to define the sequence requirements for high frequency transposition of Tn7 to the chromosomal attTn7 sites. One approach was by deletion and saturated point mutagenesis studies; therefore, it was important to obtain attTn7 site mutants in a vector on which the transposition assay could be performed directly.

Previous work of Lichtenstein and Brenner (1982) has shown that a 958 bp fragment which encompasses the E. coli attTn7 site serves as an efficient target site for Tn7 transposition when cloned in a high copy number plasmid. This observation was exploited to create a transposition assay system in which attTn7 resides in M13 phage vectors mp8 in one orientation and mp10 in the other orientation with reference to the priming site of M13, mp series vectors. The DNA sequence analysis shows that Tn7 inserts 86 bp inward from one end of the 958 bp fragment encompassing the attTn7 site (Lichtenstein and Brenner, 1982).

The transposition of Tn7 to M13 phages containing the attTn7 can be deduced by direct measurement of trimethoprim resistant (Tn7 encoded) transductants which results from Tn7 transposition. The phages which are propagated on a bacterial strain containing Tn7 in the chromosome or in a plasmid can be scored for trimethoprim (Tp) resistance on a trimethoprim sensitive recipient bacterial strain. The number of Tp resistant transduced phages can be compared with the total plaque forming unit (PFU) of the phage stock which can be expressed as a ratio which is transduction frequency for each attTn7 site. A random "generic brand" non-attTn7 DNA can be used to make a comparison between the high frequency attTn7 site transposition versus the cold site transposition.

3.2 Advantages of Using M13 Phages:

M13 phages used for the studies on Tn7 were adopted due to several reasons but primarily for the fact that they can be used as vectors for DNA sequencing of the mutant attTn7 sites by the Sanger dideoxy chain termination method (Sanger et al., 1977).

M13 is a single-stranded filamentous male-specific E. coli bacteriophage. Upon infection single-stranded phage DNA is converted into double-stranded DNA (RF form) as an intermediate in the phage life cycle. The ease with which single-stranded and RF DNA of M13 phage can be obtained is

a pronounced advantage for in vitro manipulation of DNA. Bacterial growth is slightly retarded by M13 infection and therefore infected bacteria can be detected by turbid "plaque" formation. To allow an easy selection for cloning in the M13 phages, M13, mp series vectors were constructed by Messing and coworkers, by inserting the lac regulatory region and the alpha peptide of beta-galactosidase in M13 RF DNA in vitro (Messing et al., 1977). This facilitates the identification of clones in M13 phage vectors in the presence of isopropyl-beta-D-thiogalactopyranoside (IPTG, an inducer for lacZ) and 5-bromo-4-chloro-3-indolyl-beta-galactopyranoside (X-gal, a colourless analogue of lactose) on a bacterial lawn with deletion in the alpha peptide of beta-galactosidase.

The intra-cistronic complementation of alpha and beta peptides of beta-galactosidase yields a functional activity which hydrolyses the lactose into galactose and glucose. X-gal is hydrolysed by beta-galactosidase to release the deep blue and insoluble 5-bromo-4-chloro indole. Any insertion in the lacZ region of M13 phage vectors will inactivate the beta-galactosidase, therefore, the blue morphology of the plaques will not be observed in recombinant M13 phages. However, in rare cases translational read-through of the insert can still produce blue plaques upon selection on plates containing X-gal and IPTG. Although this selection procedure was not extensively needed in this system, it was hoped to clone

the minimum attTn7 sequence "in frame" in beta-galactosidase of M13 phage vectors. This may have allowed for another system to measure the Tn7 transposition on a bacterial lawn containing Tn7, simply by scoring the white versus blue M13 plaques on plates containing X-gal and IPTG. The white plaques would have been generated upon transposition of Tn7 in the "in frame" minimum attTn7 sequence.

Since M13 phage particles contain the single-stranded DNA, this form of DNA can be used as a target for introducing base substitution in the cloned DNA (Myers et al., 1985; Smith, 1982; Zakour and Loeb, 1982; Zoller and Smith, 1983). These methods can be used to generate large numbers of independent point mutants in the attTn7 cloned in M13 phage vectors and the mutants with an altered transduction frequency can be identified easily.

Filamentous single-stranded DNA phages produced by an infected male strain of E. coli have been shown to transduce kanamycin and ampicillin resistance (Cohen, 1976). This transduction was due to the integration of transposons (confering resistance to kanamycin and ampicillin) into a site in the intergenic space between genes II and IV of the M13 genome, thus expressing the antibiotic resistance genes.

In M13 phage all genes are necessary for phage growth, so it would be expected that most transducing phages (with

transposition of Tn7 into these genes) would yield defective phages. However, high frequency, site-specific transposition of Tn7 into the attTn7 site (a 958 bp fragment from the E. coli chromosome) allowed the development of a novel phage with the attTn7 site cloned into M13 phage vector, into which Tn7 was expected to transpose specifically. This high specificity of transposition of Tn7, which has been shown to be dependent on the tnsD protein of Tn7, could be determined by scoring the trimethoprim resistance on M13 phages containing the "hot site" versus the M13 phages containing a "cold site". In the latter case it could be a random "generic brand" non-attTn7 DNA (i.e. which is not a target site for Tn7 transposition) in which case the transposition is tnsE dependent. The tnsD dependent reaction for the transposition of Tn7 in the "hot site" is higher by 3-4 orders of magnitude than the tnsE dependent in the "cold site" (Rogers et al., 1986). Therefore, this value could be compared easily in M13 phages containing separate attTn7 and non-attTn7 sites DNA.

3.3 Factors Affecting the Transposition of Tn7 into M13 Phages:

Since Tn7 is a very large transposon, a major concern in this assay was the problem of the stability of the M13 genome after transposition of Tn7. The factors which could

have some effect on the measurement of transducing phages by this method are discussed in sections 3.3.1 and 3.3.2.

3.3.1 Excision and Deletion Associated with Transposition of Tn7 in M13 Phages:

In the assay system a recA⁺ bacterial strain 71/18 was used as the donor for Tn7, therefore recA promoted DNA rearrangements were expected in M13 phages containing Tn7. The intra-/inter-molecular recombination between the inserted Tn7 in M13 genome, may generate deletion products of variable size.

By the application of a Poisson distribution, a proportional relationship can be calculated between the transposition of Tn7 and the associated effects of transposition. The measure of a transposition event (T) in M13 phage is directly proportional to any intrinsic value of deletion or excision (DE) of M13::Tn7.

$$T \propto DE \quad \text{or}$$

$$T = K (DE)$$

In this formula, K has a constant value, therefore any effect of the phenomenon (DE) may not be significant in the assay as it will occur uniformly regardless of which attTn7 site is used. Nevertheless, the overall deletion and excision events occur at a low frequency in comparison to the actual transposition events into plasmids or phages in E. coli (Kleckner, 1981).

Single-stranded DNA may also permit precise excision of Tn7 from the M13 recombinant clones. The F' factor of E. coli has been shown to stimulate the precise excision of Tn5 and with low frequencies in a recA host (Berg et al. 1983a). They showed that the ability of F' to stimulate precise excision is dependent on the conjugal transfer of F' factor. A model for transposase independent excision of Tn5 was proposed by Berg and Berg (1983). In this model, Tn5 excision was predicted in a fixed manner by denaturation of both strands of the element, followed by reannealing at the ends. The continuation of the chain growth (DNA synthesis) proceeds through the junction, thus failing to copy Tn5 during DNA synthesis of the replicon on which Tn5 resides. This model can also apply to the F' transfer which occurs at a single-stranded level. In this case, the intra-molecular reannealing between two complementary sequences in the inverted repeats of transposable elements would allow a pause for the DNA polymerase during DNA replication leading to a precise excision of the transposons.

This type of DNA rearrangement may occur in this assay system, as the M13 life cycle involves the synthesis of single-stranded DNA and thereby can allow excision of Tn7 from M13 genome. The phenomenon of this class would again precisely excise the Tn7 uniformly from M13 phages by error-prone DNA synthesis, regardless of which attTn7 site is used in the assay. Therefore, this factor does not

interfere with the correlation between transposition and transduction frequencies.

However, Tn7 transposition into M13 phage vectors may make the task of packaging the phage particles more difficult. Moreover, deletions may be obtained due to the concomitant instability of the genome. The duplication of sequences may also occur as a result of recombination between viral genomes carrying identical transposable elements inserted at different sites.

Above all, the deletion and excision associated with phage replication or Tn7 transposition will occur uniformly. Therefore, it is reasonable to believe that the background of both phenomena would not cause any noticeable difference on the validity of this assay.

3.3.2 Time Course for Transposition of Tn7:

Replication of M13 DNA proceeds through three consecutive stages: (I) conversion of viral single-stranded DNA to the double-stranded replicative form ($SS \rightarrow RF$), (II) multiplication of replicative forms ($RF \rightarrow RF$) and (III) synthesis of progeny single-strand ($RF \rightarrow SS$). Stages II and III employ M13 duplex DNA as template and depend on phage gene II product. The switch from double-strand to single-strand is brought about by the accumulation of M13 gene IV product (see review by Ray, 1978).

It is during stage II, that the transposition of Tn7 into the M13 phage vector is presumed to take place. The transposition of Tn7 may involve single-stranded intermediates but is not likely to take place in single-stranded DNA, as the 5 bp duplication in the target site would not be ensured by this mode of transposition.

During M13 phage propagation, the timing of the switch from stage I to II and from II to III in relation to the timing of a Tn7 transposition event could have some bearing on the estimation of the transduction frequency. Furthermore, the product of the initial transposition event in M13 phage can be released from the bacterial cells and can go through multiple rounds of infection and replication. This would amplify the initial Tn7 transposition events and could result in fluctuations in transposition frequency from one infected cell to another. This would again happen uniformly, regardless of which M13 phages are propagated in the Tn7 donor bacterial strain.

Despite the possible factors, mentioned above, which might affect the measurement of Tn7 transposition frequency a direct selection of the transducing phages encoding trimethoprim resistance in a trimethoprim sensitive background allowed the development of this assay. Thus the relative frequency of transposition of Tn7 into M13 phage containing the attTn7 site can be inferred by direct measurement of the number of transducing phage versus the

total number of PFU of the phage stock. This ratio for a given target attTn7 can be expressed as follows;

$$\text{Transduction frequency} = \frac{\text{Total number of trimethoprim resistant transduced colonies}}{\text{Total number of PFU}}$$

3.4 Development of the Transduction Assay in M13 Phages Containing E.coli attTn7 Site:

The M13 transduction assay provided a mean to test the points mentioned above concerning the stability of M13::Tn7 phage particles and later allowed the use of this assay in addressing questions regarding the requirement of the Tn7 target site.

The assay was performed on M13 phage recombinants which contain a 958 bp fragment (flanked by EcoRI linkers) encompassing the E. coli attTn7 site. This fragment was cloned in one or other orientation into the EcoRI RE sites of M13 vectors M13, mp8 (giving VHV12) and M13, mp10 (giving VHV22) respectively. The 958 bp attTn7 fragment had previously been shown to retain function on a plasmid vector with respect to high frequency site-specific transposition of Tn7. The control in this experiment was an M13, mp8 vector containing a random 800 bp fragment (non-attTn7) from the Bacillus subtilis chromosome.

The E. coli K12 strain 71/18 [lac⁻, pro⁻, SuIII⁺, thi⁻,

F'proA⁺ +B⁺, lacI, Z^q, deletion M15] harbouring the plasmids F' (pro⁺) and the high copy Tn7 donor colE1::Tn7 was used for propagation of M13 phages. As the F' factor carries a proA⁺ +B⁺ and it complements a deletion in the chromosome, the stocks of this strain were maintained on minimal plates supplemented with 1% glucose and 10µg/ml of thiamine to select for proline prototrophy and hence retention of the F' factor.

The M13::attTn7 and control phage recombinants were used to infect an early log phase culture (at OD₅₅₀ = 0.2 units with an m.o.i. = 50) of 71/18 (ColE1::Tn7). After 4-5 hours of growth the phages ensuing from this infection were used to measure the total number of PFU of a phage stock.

Serial dilutions of each phage stock were then used to infect either a bacterial strain JM101 (pEAL1) or TG2 (at OD₅₅₀ = 0.6 units) for 10-15 minutes at 37 C for phage adsorption prior to antibiotic selection. Bacterial strain JM101 contains a plasmid pEAL1 (Tc^r) and TG2 contains a transposable element Tn10 (Tc^r) in the chromosome, and therefore in both cases tetracycline can be used as a counter selection against the donor 71/18 strain. The trimethoprim resistant transduced colonies were scored on Mueller Hinton plates containing 100µg/ml trimethoprim and 10µg/ml tetracycline. A diagrammatic representation of M13 transduction assay is shown in figure 3.1. The results obtained from this assay are summarised in table 3.1.

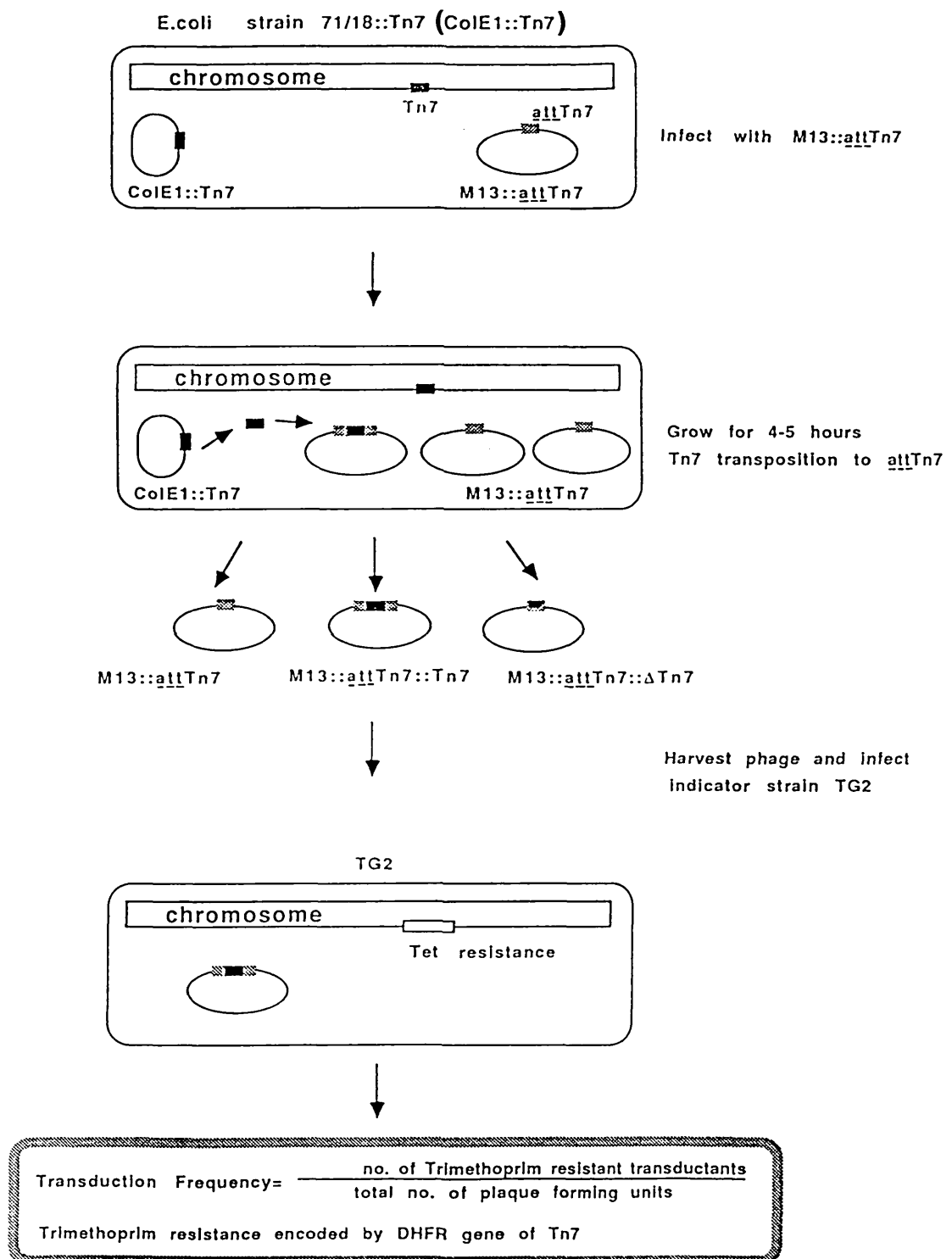


Figure 3.1 Diagrammatic representation of M13 transduction frequency assay.

VHV12 and VHV22 phage vectors have very high transduction frequencies with the range of $1\text{E}-3$ to $5.5\text{E}-4$ in comparison to the control $1\text{E}-10$ in the recipient bacterial strain JM101 containing pEAL1. Therefore, a difference of approximately one million fold was observed. This assay was performed several times and the transduction frequency results were found to be consistent. The results obtained for attTn7 site were also of several orders of magnitude higher than the nonattTn7 site in the recipient bacterial strain TG2. However these results were variable and of lower frequency with the range of $1\text{E}-4$ to $1\text{E}-5$. After repeating the assay in both recipient bacterial strains several times, a variation (about 5 to 10 fold) was found consistent. Therefore the chances for any quantitative error were less.

Two conclusions could be drawn from this analysis. First, the M13 transduction assay for the measurement of transposition frequency of Tn7 proved to work and there is significant difference (about one million fold) in the calculated transduction frequency between phages containing the attTn7 site and "generic brand" (non-attTn7) DNA. Second, the transduction frequency measurements fall within 5 to 10 fold range for a given pair of host and recipient E. coli strain and a given phage. The variation in the transduction frequency of both controls (VHV12 and VHV22) when plated on the two recipient strains, may be due to the followings reasons.

The higher number of trimethoprim resistant transductants in JM101 may be due to recA promoted recombination events. The recA protein may be acting during some stage after transposition in the recipient cells which may include a second transposition event from M13 phages containing intact Tn7, to the attTn7 site on the plasmid pEAL1 or into the chromosome. In another experiment, the M13::attTn7 phage particles were propagated in the strain 71/18 (ColE1::Tn7) and plated on either recipient strain JM101 or JM101 (pEAL1) and TG2. The transduction frequency in the recipient recA⁺ cells (either JM101 or JM101 containing pEAL1), shows only a 2 fold increase in the measured transduction frequency in comparison to recA⁻ cells TG2. The result suggested that neither the recA phenotype nor the the presence of an attTn7 containing plasmid in the recipient has a significant effect. However, it was shown in this laboratory that Tn7 transposition into M13 phage in a recA⁺ Tn7 donor host is of 10 to 50 fold higher than in a recA⁻ host (Hulton, 1987).

Another possible source of variation in the calculated transduction frequency may be due to slight variation in the incubation time prior to the antibiotic selection. However, upon infection with M13 phage into F' bacterial cells, no mature phage particles are released from the cell for 30 minutes.

The factors described above and as well as the physiological state of the donor cells at the time of infection may have a combined effect that results in the variation of the calculated transduction frequency.

3.5 M13 Transduction Assay on Junction attTn7 Sites (Tn7-attTn7 Fragment):

By employing the strategy described in section 3.4, it was possible to test a set of Tn7-attTn7 junction sites as targets for Tn7 transposition. These attTn7 sites, which contain the left and right junctions between the ends of Tn7 and the E. coli chromosome, were constructed by Conrad Lichtenstein (1980). M13 phage vector 31, 34, Pst9, Pst10, OE6 and 1B3F were employed in the M13 transduction assay. The sizes of the junction fragments containing the right or left end of Tn7 and right or left end of attTn7 are variable. 31, 34, Pst9 and Pst10 all contain a 389 bp of the right end of attTn7 from E. coli chromosome. 31 and 34 have only 41 bp of the right end of Tn7 while Pst9 and Pst10 contain an additional 495 bp of Tn7. The left end junction fragment clones, OE6 and 1B3F contain 102 bp and 133 bp of the left end of attTn7 site from E. coli chromosome and from plasmid ColE1 respectively; in conjunction with the 125 bp and 105 bp of left end of Tn7 respectively. The orientation of each junction fragment clone with reference to the priming site of M13, mp series vectors is shown in figure 3.2.

The transduction frequency results of each junction fragment (Tn7-attTn7) are shown in table 3.1 (A). In order to overcome the variation in the calculated transduction frequency, the data obtained for each tested attTn7 site is compared with that of controls (VHV12 or VHV22) to obtain the normalised transduction frequency values. These relative values of each junction fragment are shown in table 3.1 (B). The mean of the normalised transposition frequency values are indicated in the text for comparative purposes.

The transduction frequency results obtained from two sets of recipient bacterial strains, JM101 (pEAL1) and TG2, are shown in graphical form (see figure 3.2). These graphs were plotted with modified transduction frequency values as described in the figure legend.

The results obtained from this analysis are summarised in figure 3.3 and can be rationalised as follows.

The most important information necessary for high frequency transposition of Tn7 seems to reside within the right end of the attTn7 site (phage vectors 31, 34), and the left end of the attTn7 site (phage vector 0E6) appears not to function on its own. The transduction assay performed on VHV12 and VHV22 had previously suggested that the 86 bp of the right end of the attTn7 site were participating in the high frequency transposition (see figure 3.3).

Table 3.1 Results of the M13 transduction frequency assays.

A: The assays were performed using 71/18 (ColE::Tn7) as a Tn7 donor bacterial strain and JM101 (pEAL1) and TG2 as recipient bacterial strain for scoring trimethoprim resistant transduced colonies.

Target	Independent trials							
	Recipient strain JM101 (pEAL1)				Recipient strain TG2			
	1	2	3	4	1	2	3	
VHV12	5.6E-4	1.0E-3	2.8E-4	5.5E-4	5.0E-5	1.8E-4	1.5E-4	
VHV22	3.8E-4	1.5E-3	2.5E-4	5.1E-4	6.5E-5	1.0E-4	1.1E-4	
VCV8	1.0E-10	4.0E-10	4.0E-11	3.0E-10	1.0E-10	3.0E-10	3.0E-10	
31	3.5E-5	6.2E-5	1.5E-5	4.3E-5	4.5E-6	1.2E-5	1.2E-5	
34	4.0E-5	9.5E-5	2.3E-5	3.5E-5	4.8E-6	1.1E-5	8.4E-5	
Pst9	5.0E-7	1.0E-6	5.5E-7	1.6E-6	2.6E-8	4.0E-7	3.8E-7	
Pst10	1.5E-6	1.6E-6	3.5E-7	6.5E-7	1.2E-8	8.1E-8	3.5E-7	
OE6	3.0E-10	4.0E-10	2.0E-11	2.0E-10	1.0E-10	4.0E-10	3.0E-10	
1B3F	1.0E-10	3.0E-10	4.0E-11	4.0E-10	1.0E-10	3.0E-10	2.0E-10	

B: Relative values of transduction frequencies of the tested clones versus that of controls VHV12 or VHV22. The upper value is relative to that of VHV12 and the lower to that of VHV22.

Target	Independent trials									
	Recipient strain JM101 (pEAL1)					Recipient strain TG2				
	1	2	3	4	Mean	1	2	3	Mean	
VCV8	1.7E-7	4.0E-7	1.4E-7	5.4E-7	3.1E-7	2.0E-6	1.6E-6	2.0E-6	1.8E-6	
	2.6E-7	2.6E-7	1.6E-7	5.8E-7	3.1E-7	1.5E-6	3.0E-6	2.7E-6	2.4E-6	
31	6.2E-2	6.2E-2	5.3E-2	7.8E-2	6.3E-2	9.0E-2	6.6E-2	8.0E-2	7.8E-2	
	9.2E-2	4.1E-2	6.0E-2	8.4E-2	6.9E-2	6.9E-2	1.2E-1	1.0E-1	9.6E-2	
34	7.1E-2	9.5E-2	8.2E-2	6.3E-2	7.7E-2	9.6E-2	6.1E-2	5.6E-2	7.1E-2	
	1.0E-1	6.3E-2	9.2E-2	6.8E-2	8.0E-2	7.3E-2	1.1E-1	7.6E-2	8.6E-2	
Pst9	8.9E-4	1.0E-3	1.9E-3	2.9E-3	3.6E-3	5.2E-4	2.2E-3	2.5E-3	1.7E-3	
	3.9E-3	6.6E-4	2.2E-3	3.1E-3	2.4E-3	4.0E-4	4.0E-3	3.4E-3	2.6E-3	
Pst10	2.6E-3	1.6E-3	1.2E-3	1.1E-3	1.6E-3	2.4E-4	4.5E-4	2.3E-3	9.9E-4	
	3.9E-3	1.0E-3	1.4E-3	1.2E-3	1.8E-3	1.8E-4	8.1E-4	3.1E-3	1.3E-3	
OE6	5.3E-7	4.0E-7	7.1E-8	3.6E-7	3.4E-7	2.0E-6	2.2E-6	2.0E-6	2.0E-6	
	7.8E-7	2.6E-7	8.0E-8	3.9E-7	3.7E-7	1.5E-6	4.0E-6	2.7E-6	2.7E-6	
1B3F	1.7E-7	3.0E-7	1.4E-7	7.2E-7	3.3E-7	2.0E-6	1.6E-6	1.3E-6	1.6E-6	
	2.6E-7	2.0E-7	1.6E-7	7.8E-7	3.5E-7	1.5E-6	3.0E-6	1.8E-6	2.1E-6	

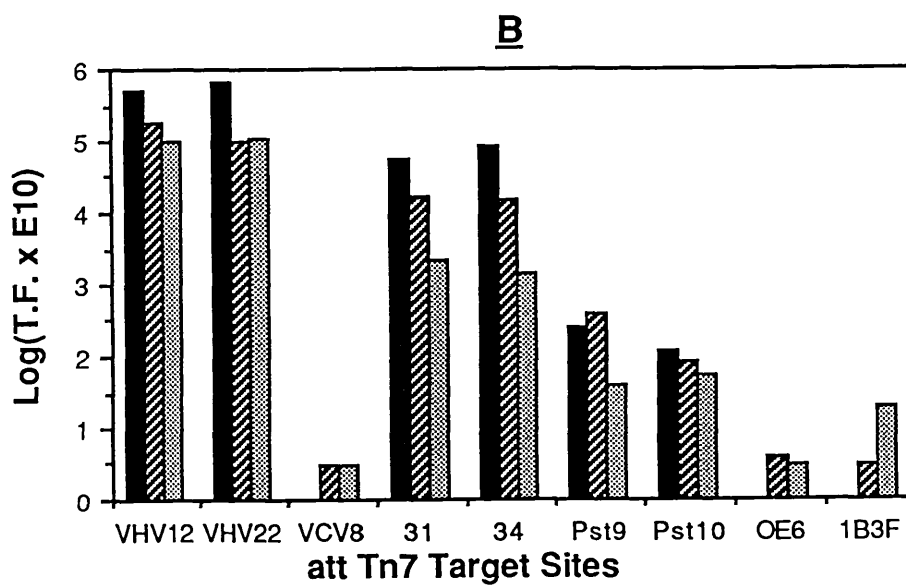
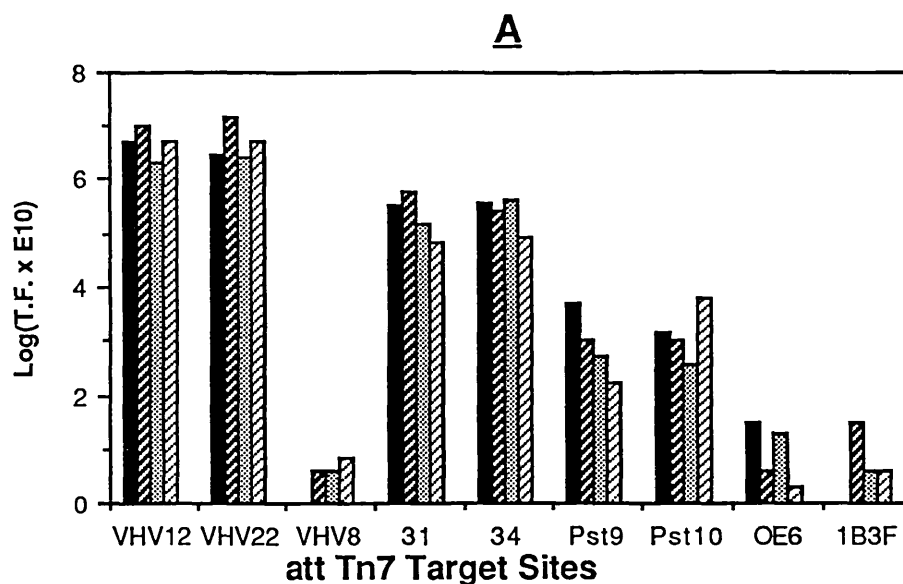


Figure 3.2 Graphical representation of the results of M13 transduction frequency assays on different attTn7 sites. In order to plot the data, the values for the transduction frequency were scaled-up to more workable numbers. This was done by multiplying each set of actual transduction frequencies by an appropriate power (indicated in each diagram), and then by plotting the logarithms of these scaled-up values against attTn7 target sites.

A: graph of results obtained with the recipient bacterial strain JM101 (pEAL1).

B: graph of results obtained with the recipient bacterial strain TG2.

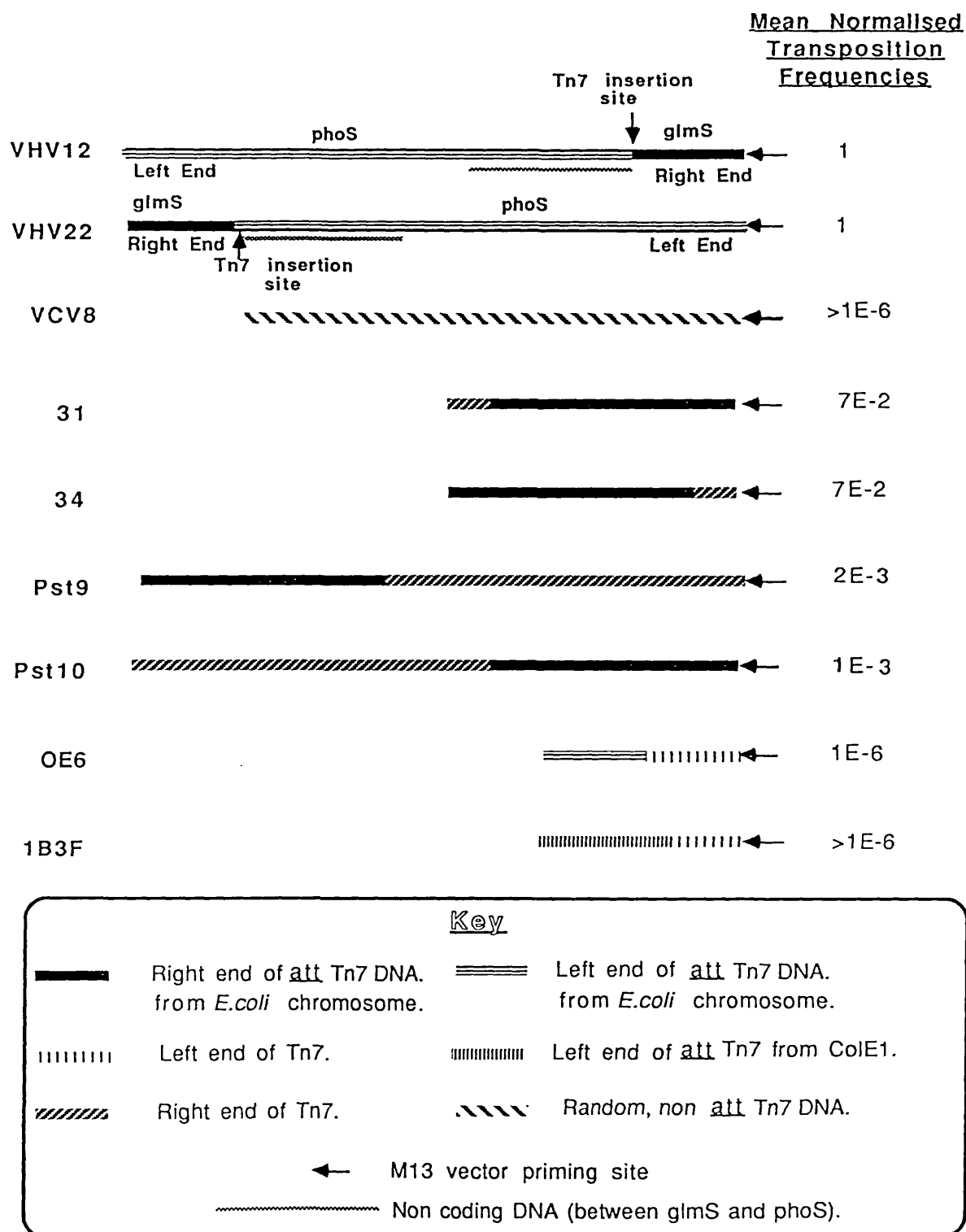


Figure 3.3 Summary of the results of M13 transduction frequency assay of different attTn7 sites.

A drop of 10 to 15 and 11 to 14 fold in the transduction frequency was observed while using the M13 clones 31 and 34, respectively. This may be due to the removal of some essential DNA sequences in attTn7 site upon insertion of Tn7 or this insertion may have destroyed the transcriptional terminator of the glmS which may be required for high frequency transposition of Tn7.

However a simple conclusion can be drawn from this result that the 86 bp of the right end of attTn7 site can provide an efficient target for Tn7 transposition whilst the left end of attTn7 can not. Therefore, the right (glmS) end of the attTn7 site seems to be more important than the left (phoS) end for target site selection in high frequency transposition of Tn7 in the cloned 958 bp attTn7 site.

A drop of 384 to 625 and 555 to 1000 fold was observed in the transduction frequency of Tn7 in the right end junction attTn7 sites Pst9 and Pst10, respectively. This difference of the calculated transduction frequency in comparison with 31 and 34 was significant with the fact that all four M13 clones contain the identical attTn7 sequence but variable length Tn7 sequence. How can an additional 495 bp of the right end of the Tn7 negatively regulate the transposition of Tn7? This can be explained on the following basis.

(i) A regulatory sequence in the right end of Tn7 may be responsible for providing binding sites of Tn7 transposase, thus competing with the ends of Tn7 on the

donor replicon.

(ii) A repressor activity of the truncated tnsA protein encoded on the 536 bp of the right end of Tn7 may be interfering with the function of the transposase of Tn7.

(iii) The presence of the 536 bp right end of Tn7 is substantially reducing the frequency of subsequent transposition by a phenomenon referred to as transpositional immunity. This phenomenon has been observed in the cases of Tn3, phage Mu and possibly Tn554 (see chapter 1, section 1.3.1).

Hauer and Shapiro (1984) also believe that Tn7 displays transpositional immunity. Therefore transpositional immunity provided by the right end of Tn7 may be the cause that the attTn7 sequence in atttTn7::rightTn7⁵³⁶ does not provide an active target for subsequent Tn7 insertion.

A conclusion drawn from this analysis could be that the 41 bp of the right end of Tn7 are not sufficient for this transpositional inhibition; the DNA sequence within the 536 bp of the right end of Tn7 contains all necessary information for this phenomenon. To define the extent of DNA sequence required within the 536 bp of the right end of Tn7 for transpositional immunity, unidirectional deletion analysis was required to be performed. However, this approach was not taken.

3.6 M13 Transduction Frequency Assay of Tn7 Utilising Different Bacterial Host Strain:

The M13 transduction assay can be exploited to probe the involvement of any host encoded functions either in the regulation or mechanism of Tn7 transposition. The requirement for host encoded function DNA gyrase was shown for transposition of Tn5 (Isberg and Syvanen, 1982). They observed that if the DNA acting as a recipient of the transposon is deficient in supercoils, it is a poor substrate for transposition. Therefore, the functional DNA gyrase is necessary for the supercoiling of DNA.

A variety of E. coli mutants have been isolated in the genes for DNA processing (e.g. gyrase A, gyrase B, topoisomerase), genes involved in DNA rearrangement (e.g. recA, recBC, recF) and genes for transcriptional control (e.g. nusA, nusB). Some of these mutants can be used to find the dependency of Tn7 transposition on their gene product by using the M13 transduction frequency assay.

The mutations in nusA and nusB prevent pN (protein N) function from interacting with the beta sub-unit of the core RNA polymerase. pN has been shown to act as an antiterminator protein which binds with the nut site on bacteriophage lambda. Antitermination mechanism has been shown to be a positive control of transcription for the expression of gene products for bacteriophage lambda growth (Freidman et al., 1984) and for the ribosomal RNA operon of E. coli (Morgan, 1986).

The transposition of Tn7 may involve such antitermination regulations involving a host protein like pN. Smith and Jones (1986) showed that tnsE gene product is involved in the post-transcriptional regulation of genes downstream of tnsB by preventing the transcriptional termination between the HindIII RE site at position 3042 and the BamHI RE site at position 3937 in the right end of Tn7; a tnsD encoding fragment. Whether or not it is one positive control for Tn7 transposition remains unclear. One way to answer this question can be the use of a bacterial strain in the M13 transduction frequency assay which have mutations in nusA or nusB. Mutations in these genes would prevent the tnsE protein of Tn7 from interacting with the RNA polymerase at the putative transcriptional terminator and thereby decreasing the transposition frequency.

In this study a host strain with mutations in nusA was employed to measure the transduction frequency and the strategy is discussed below.

Tn7 was delivered to bacterial strain LP144 (nusA⁺, Tcr) and LP145 (nusA⁻, Tc^r) through the F' episome from a bacterial strain harbouring ColE1::Tn7. 71/18 (ColE1::Tn7) was naturally selected for the loss of ColE1 while F' and Tn7 were maintained on minimal glucose plates containing 100µg/ml trimethoprim. The Colicin E1 sensitive 71/18 colonies were tested for the loss of ColE1 as described in chapter 2, section 2.2.1(iii).

The transfer of F':::Tn7 to LP144 and LP145 was achieved via an intermediate host Δ7. The bacterial strain (Δ7) was constructed by Carlos Flores in this laboratory and the strategy is described in chapter 5, section 5.7. The reasons to use the Δ7 as a transient host were two fold; (1) this could provide an easy selection procedure for the delivery of F':::Tn7 to LP144 and LP145 as it has glmS phenotype, and (2) Δ7 could be tested for the presence of any secondary insertion site of Tn7 in the chromosome as the primary attTn7 was deleted from the chromosome of this strain. The second reason is addressed in chapter 5, section 5.7. The transfer of F':::Tn7 in bacterial strain Δ7 was achieved by mating the recipient and donor [71/18, (F':::Tn7)] colonies on MH plates. The exconjugants were then selected on MH plates containing Kn, Tp and 10μg/ml glucosamine. Since the plasmid ColE1 was lost in the Tn7 donor strain 71/18, the Tn7 encoded antibiotics can be expressed only following the successful transfer of F':::Tn7.

After ensuring the transfer of F':::Tn7 into bacterial strain Δ7, 100 bacterial colonies were tooth-picked in duplicate onto MH plates and each colony was mixed with the cells of LP144 and LP145 for the delivery of F':::Tn7. The exconjugants of LP144 and LP145 were selected on MH plates (without glucosamine) supplemented with Tp and Tp plus Tc^r respectively. Tetracycline resistance gene is expressed only in LP145 due to the presence of Tn10 (Tc^r) in the

nusA gene. Three recipient colonies each for LP144 (F':::Tn7) and LP145 (F':::Tn7) were picked and grown in 2X YT media which were then infected with M13 phages containing attTn7 DNA (VHV12 and VHV22) and non-attTn7 DNA (VCV8). The M13 transduction assay was performed on these bacterial strains and an approximately 1000 fold drop was observed in all three cases of LP145 (F':::Tn7) when compared to the isogenic strain LP144.

The F' could not be maintained in LP145 as the host strain has pro⁺ genotype, therefore this preliminary observation may need to be repeated. However, this result has suggested an interesting speculation about the involvement of host protein nusA in the antitermination regulation of Tn7 transposition as proposed by Smith and Jones (1986).

The choice of Tn7 donor replicon did not influence the transduction frequency. This was observed when the Tn7 was either in the chromosome, i.e. 71/18::Tn7 (constructed in this study) or in the high copy number plasmid, i.e. pUC18attTn7::Tn7 (pFK::Tn7).

CHAPTER 4

Definition of the Minimal Target DNA Sequence for High Frequency Transposition of Tn7 to the E. coli attTn7 Site

4.1 Introduction:

In this study, the objective was to define a minimum attTn7 DNA sequence which can function as an efficient target site for the transposition of Tn7. The target for this analysis was a 958 bp fragment encompassing attTn7 which was cloned in the M13 phage vectors, mp8 and mp10 at the EcoRI site in both possible orientations with reference to the universal priming site of the M13, mp series.

The approach taken to meet this objective was to delete the DNA from the right and the left end of the 958 bp attTn7 cloned into these M13 phage vectors and to test the effect of these deletions on the transposition frequency of Tn7 using the M13 transduction assay as described in chapter 3.

By doing this it was possible to locate the extent of DNA sequence required for high frequency, site-specific transposition of Tn7.

4.2 Deletion Mutagenesis:

A variety of methods can be used to obtain variable deletions at the ends of the DNA sequence. One method uses the Bal31 exonuclease. This enzyme can degrade both single-stranded and double-stranded DNA in both 5' to 3' and 3' to 5' directions. Thus treatment of linear double-stranded DNA with this enzyme, followed by ligation of the blunt ends with T4 DNA ligase, generates circular plasmid DNA with deletions of various lengths. This approach was used to analyse the transcriptional regulatory signals of the Herpes Simplex Virus I, thymidine kinase gene (McKnight et al., 1981). This method was not suitable for the analysis of attTn7 required here, as deletions at both ends would have produced defective M13 phage particles and required that the attTn7 region be recloned.

Bal31 generated deletions can also be introduced by methods that generate random nicks in duplex DNA. For example, Pancreatic DNase I, in the presence of manganese ions, makes random double-stranded breaks in DNA. Partial digestion of duplex circular plasmid or viral DNA in this way can generate circularly-permuted full length linear molecules, which can serve as a substrate for Bal31 to generate deletions. Whilst this method will produce internal deletions which could include deletions at one or other end of the target DNA sequence, it is not end specific per se.

An in vivo method to generate end-specific deletions involves the use of transposable elements. This method exploits a property of transposable elements, associated with transposition, namely generation of deletions from one or other end of the element (see chapter 1, section 1.2).

Transposon Tn9 has been used in this way for the functional analyses of cloned yeast TRP1 and ARS1 sequences in a vector containing Tn9 and part of the bacterial gal operon, i.e. galK and T (Ahmed, 1984). This vector confers Ap^r, trp⁺, Tc^r and Cm^r phenotype and is used in conjunction with the E. coli strain, Δ4 [F⁻, recA, trpC, 4(gal-chlD-pgl-attlambda) sup^o, strA].

Deletions start from one fixed site at the right terminus of the IS1 insertion element of Tn9. These deletions are presumed to result from the resolution of cointegrates formed by intra-molecular transposition. Thus, a fixed site at the transposon terminus is fused to variable sites on adjoining DNA in different deletion events. If a transposon with terminus A is located next to DNA of the sequence BCDE, then the deletion events would generate DNA segments of the type ABCDE, ACDE, ADE and AE. These deletions can be isolated by positive selection for galactose resistance (acquired by deletion of the galK gene) by plating the bacterial cells containing a recombinant plasmid on MacConkey-galactose plates. Thus, all gal^r colonies which appear after selection, are the

product of deletions extending from the right terminus of Tn9 to various points in the galK gene beyond, i.e. into the cloned insert of itself.

Therefore, using a primer complementary to A, the nucleotide sequence of each deletion mutant can be determined providing an overlapping set of DNA sequences. This elegant approach was not applicable to analysis of attTn7 as it would have involved a laborious recloning of each deletion mutant from the plasmid into M13 phage vectors to use in the M13 transduction assay. The possibility of using the Tn9 based deletion system directly in M13 phage vectors containing attTn7 was not chosen as it would have required the development of a new M13-based vector.

The method adopted here was to obtain unidirectional deletions of the attTn7 DNA by exploiting the property of exonuclease III. Exo III is a double-stranded exonuclease which digests in the 3' to 5' direction. However, it requires a recessed 3' end (for example, the ends produced by BamHI) or a blunt end, to initiate digestion. The ends with 3' overhangs (such as those produced by PstI cleavage) can not serve as a substrate for Exo III digestion.

Thus, M13 phage vectors containing an insert can be cleaved with two types of restriction endonucleases, to yield a linear molecule with a 3' overhanging terminus towards the priming site of the M13 vector and a 3' recessed terminus

towards the cloned insert.

A 3' recessed end will be hydrolysed by exonuclease III while the 3' overhanging end will be protected from exonuclease III digestion (see figure 4.1). After Exo III treatment, S1 nuclease can be used to remove any single-stranded 3' and 5' overhanging termini from the linear molecule. T4 DNA ligase can be used to ligate these molecules in vitro. The transformants can be analysed directly by the Sanger dideoxy chain termination method for DNA sequencing (Sanger et al., 1977).

The length of deletion generated in this way can be varied by performing the Exo III digestion at varying temperatures and incubation times (Guo and Ray, 1983). The advantage of this strategy is two fold; it can provide, (i) a direct phenotypic selection of the mutants by using the M13 transduction assay and (ii) a rapid identification of the mutants by DNA sequencing (Sanger et al., 1977).

4.3 The Minimum attTn7 is in a 64 bp Fragment:

The target for unidirectional deletion analysis (the 958 bp fragment) had in its right end, a 61 bp DNA sequence of glmS encoding the carboxy-terminal domain of glucosamine synthetase and a 307 bp downstream intergenic untranslated sequence, and in its left end, 590 bp of phoS encoding the phosphate-repressible phosphate binding protein (shown previously in figures 1.10 and 3.3).

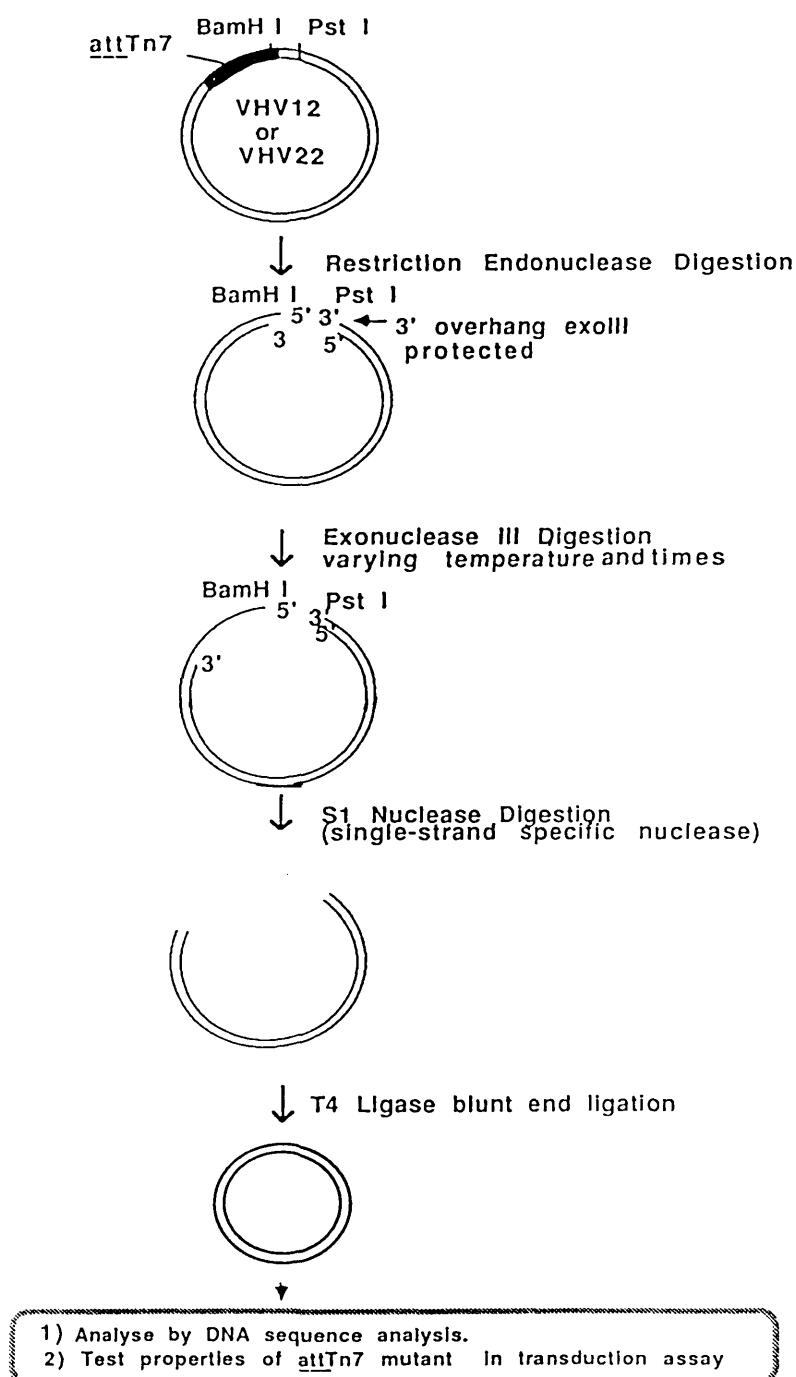


Figure 4.1 Strategy of unidirectional deletion analysis using exonuclease III and S1 nuclease.

The Tn7 insertion site was shown to be in the transcriptional terminator of the glmS gene at position 5'-GCGGG-3' between base pairs 81 and 86 of the 958 bp attTn7 (Gay et al., 1986; Lichtenstein and Brenner, 1982). This fragment was shown to be an efficient target site for Tn7 transposition when cloned into high copy number plasmids (Lichtenstein and Brenner, 1981). This attTn7 fragment was sub-cloned from pEAL1 as an EcoRI restriction fragment into the EcoRI site in M13 phage vectors, mp8 and mp10 to yield VHV12 and VHV22 respectively. In VHV12 the C-terminal region of the glmS is adjacent to the M13 universal priming site whilst in VHV22 the fragment is in the opposite orientation.

Both VHV12 and VHV22 were tested using the M13 transduction assay as described in chapter 3. The transduction frequency for both phages was found to be of the magnitude $1\text{E}-4$ to $2.1\text{E}-5$ Tp^{r} transductants per PFU, in contrast, control M13 phage containing a non-attTn7 site showed transduction frequencies of magnitude less than $2\text{E}-10$.

The DNA sequence from left (phoS) end of the attTn7 site was removed by digesting the VHV22 (RF DNA) with PstI and BamHI REs. The linear DNA was then treated with exonuclease III [see chapter 2, section 2.2.2D(vi)], removing samples at 10 minute intervals. Then after S1 nuclease treatment, the DNA molecules were ligated and transformed to

bacterial strain TG2 [see chapter 2, section 2.2.1(i)]. M13 plaques were picked at random from the plate to prepare single-stranded DNA for DNA sequence analysis; only ddTTP reactions were carried out initially.

Six deletion mutants with sequential deletions of 124, 284, 425, 693, 843 and 849 bp were identified and designated as $\Delta L124$, $\Delta L284$ etc. Figure 4.2 shows the autoradiograph of the DNA sequencing gel of longest deletion mutant $\Delta L849$. All deletion mutants were subjected to the M13 transduction assay and the results are summarised in table 4.1 (A). The relative transduction frequency values of each deletion mutant with that of control VHV22 are also expressed in the table 4.1 (B). Figure 4.3 shows the graphical representation of transduction frequency results which were plotted by the same procedure as described in figure 3.2.

It seems apparent from these results that the DNA sequence from the left end of attTn7 up to base pair 849 is not required for high frequency transposition of Tn7. The 89 bp of the attTn7 remaining in deletion mutant $\Delta L849$ appears to retain the full wild-type transduction frequency.

Further deletions extending just beyond position 849, were not obtained. This may be related to the presence of the inverted repeat. Perhaps because of the formation of a secondary structure ($\Delta G = -19.2$ Kcal per mol) of the transcriptional terminator of the glmS gene, exonuclease III proceeded very rapidly into this region.

Figure 4.2 DNA sequencing gels of deletion mutants.

A: the ddTTP track run of deletion mutants from the right end of the 958 bp attTn7 site. $\Delta R11$ to $\Delta R38$ clearly shows that there are no deletions within the 5' protected end of the vector (lacZ DNA), as the lower 2 bands in all cases are at similar positions.

$\Delta L849$: the longest deletion mutant from the left end of the attTn7 site and shows a region of 89 bp attTn7 DNA marked by arrows. The * shows the occurrence of a HincII RE site which was later utilised to obtain MAS.

MAS: the 70 bp mini attTn7 site sequence marked by arrows and is in complementary form to $\Delta L849$.

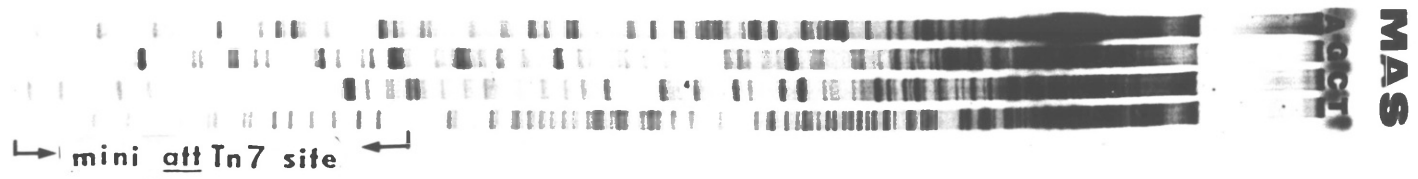
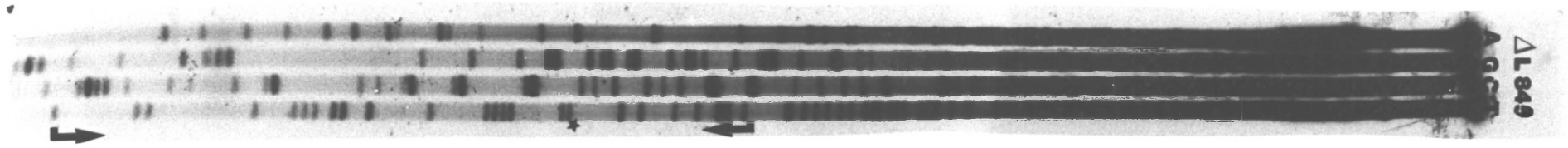
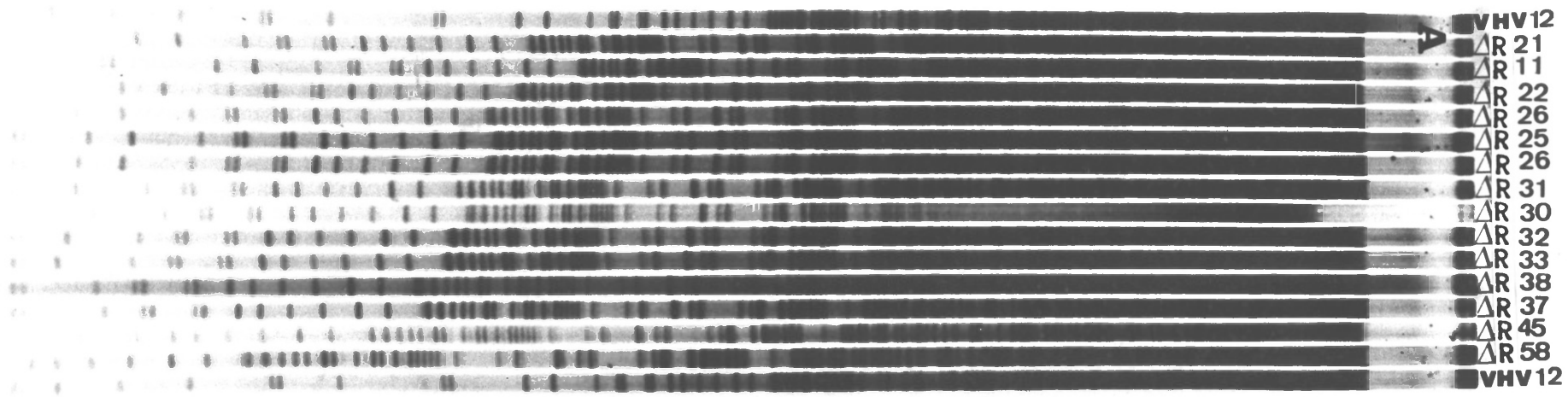


Table 4.1 Results of the M13 transduction frequency assays of deletion mutants of 958 bp attTn7 site.

A: The assays were performed using 71/18 (ColE1::Tn7) as a Tn7 donor strain and TG2 as a recipient strain for scoring trimethoprim resistant transduced colonies.

Target	Independent trials			
	1	2	3	4
VHV22	3.0E- 4	1.0E- 4	2.0E- 5	2.1E- 5
ΔL124	3.0E- 4	1.5E- 4	1.1E- 5	1.0E- 5
ΔL284	1.0E- 4	1.1E- 4	1.2E- 5	4.0E- 5
ΔL425	1.1E- 4	2.1E- 4	1.0E- 5	2.8E- 5
ΔL693	5.1E- 4	2.1E- 4	1.5E- 5	1.5E- 5
ΔL843	2.1E- 4	5.1E- 4	3.7E- 5	2.7E- 5
ΔL849	2.1E- 4	1.5E- 4	2.1E- 5	2.1E- 5
VHV12	2.0E- 4	1.3E- 4	5.0E- 5	1.0E- 5
ΔR11	1.1E- 4	2.4E- 4	5.1E- 5	2.1E- 5
ΔR21	2.1E- 4	3.4E- 5	8.0E- 6	2.1E- 5
ΔR25	1.1E- 4	1.0E- 4	3.1E- 5	1.1E- 5
ΔR26	2.1E- 5	5.4E- 5	1.5E- 5	3.8E- 6
ΔR30	1.5E- 6	9.0E- 8	4.5E- 7	1.5E- 7
ΔR31	1.1E- 6	3.4E- 8	1.4E- 7	8.4E- 8
ΔR32	1.0E-10	4.0E-10	1.0E-10	3.0E-10
to ΔR58				

B: Relative values of individual transduction frequency of the tested clones versus that of controls (VHV22 and VHV12).

Target	Independent trials				Mean relative values
	1	2	3	4	
VHV22	1.0E 0	1.0E 0	1.0E 0	1.0E 0	1.0E 0
ΔL124	1.0E 0	1.5E 0	5.5E-1	4.7E-1	8.8E-1
ΔL284	3.3E-1	1.1E 0	6.0E-1	1.9E 0	9.8E-1
ΔL425	3.6E-1	2.1E 0	5.0E-1	1.3E 0	1.0E 0
ΔL693	1.7E 0	2.1E 0	7.5E-1	7.1E-1	1.3E 0
ΔL843	7.0E-1	5.1E 0	1.8E 0	1.2E 0	2.2E 0
ΔL849	7.0E-1	1.5E 0	1.0E 0	1.0E 0	1.0E 0
VHV12	1.0E 0	1.0E 0	1.0E 0	1.0E 0	1.0E 0
ΔR11	5.5E-1	1.8E 0	1.0E 0	2.1E 0	1.3E 0
ΔR21	1.0E 0	2.6E-1	1.6E-1	2.1E 0	8.8E-1
ΔR25	5.5E-1	7.6E-1	6.2E-1	1.1E 0	7.5E-1
ΔR26	1.0E-1	4.1E-1	3.0E-1	3.8E-1	2.9E-1
ΔR30	7.5E-3	6.9E-4	9.0E-3	1.5E-2	8.0E-3
ΔR31	5.5E-3	2.6E-4	2.8E-3	8.4E-3	4.2E-3
ΔR32	5.0E-7	3.0E-6	2.0E-6	3.0E-5	9.8E-6
to ΔR58					

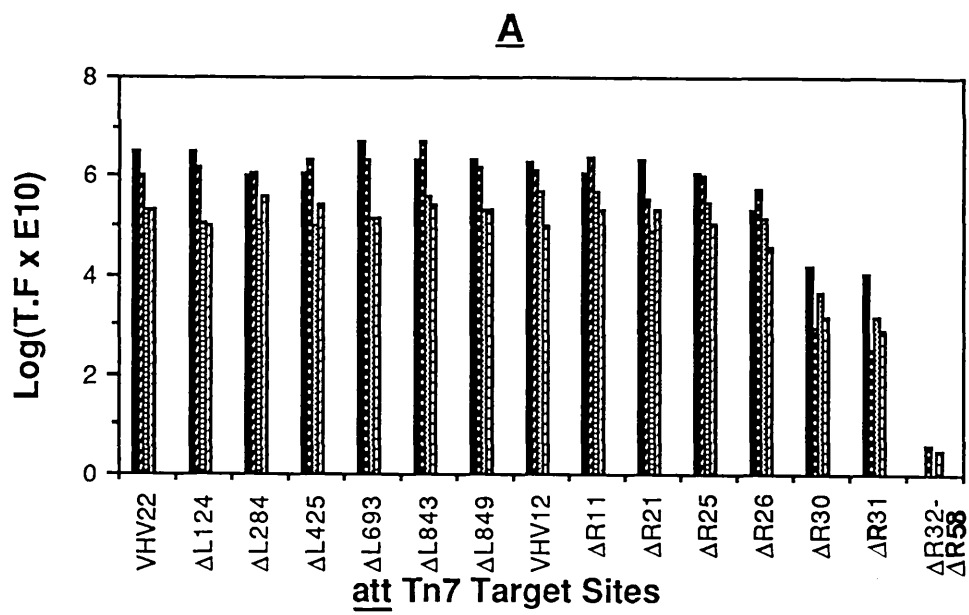


Figure 4.3 Graphical representation of the results of M13 transduction frequency assays on deletion mutants of 958 bp attTn7 site. The graph is plotted according to the same procedure as described in figure 3.3.

However, many much larger deletions were obtained extending beyond the 958 bp of attTn7 insert and into the lacI gene of the M13, mp10 vector. These, of course, were of no interest.

In order to determine the DNA sequence required from the right end of the 958 bp attTn7 fragment, deletion mutants were obtained following the same strategy of Exo III and S1 nuclease treatment. However, the DNA samples were collected after every 1 minute of Exo III treatment.

Twelve mutants were identified among the transformants of 5 minutes Exo III treatment. They were identified by DNA sequencing (ddTTP reaction only) and figure 4.2 shows the resultant autoradiograph of these mutants. In these deletion mutants, 11, 21, 25, 26, 30, 31, 32, 33, 37, 38, 45 and 58 bp were sequentially removed from the glmS coding region and designated Δ R11, Δ R21 etc. The deletion mutants were tested for transduction frequencies of Tn7 and the results are presented in table 4.1 (A) and the relative value of transduction frequency for each mutant to the wild-type (i.e. VHV12) is also shown in table 4.1 (B). Figure 4.3 shows a graphical representation of the data.

The functional importance of the DNA sequence beyond position 25 from the right end of the 958 bp attTn7 DNA was very obvious. Deletion of 25 bp has apparently no effect on the wild-type function of the target site. However, the result shows that a single bp additional

deletion from the right end of the attTn7 (Δ R26) has reduced the activity of attTn7 by approximately 3 fold in comparison to the wild-type. A further 4 bp deletion (30 bp in total) in Δ R30 reduces the transduction frequency by approximately 125 fold. An additional 1 bp deletion, (31 bp in total) in Δ R31 reduces activity by approximately 238 fold and the deletion of 1 more base pair, (32 bp in total) in Δ R32 completely abolishes the attTn7 activity, as a transduction frequency of less than $1E-10$ was observed in this deletion mutant. All other deletion mutants from the right end of attTn7 site beyond the 32 bp have more or less the same transduction frequency as that of a non-attTn7 site.

As with all deletion analysis of DNA, the effect of bringing the functional element closer to another DNA sequence (for example attTn7 site is brought closer to the lacZ or lacI DNA in M13, mp8 and mp10) can not be ignored. In the experiment discussed above this may have played some role as from the left or right end of the attTn7 a different DNA sequence was placed adjacent to deletion mutants (see figure 4.4).

In deletion mutant Δ R26, a deletion of 1 base pair in comparison with Δ R25, has brought 2 identical nucleotides of lacZ next to the glmS sequence.

3' - AATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCGACCAGTTGCAGCCACGGAAACTAG - 5'
 * wild-type

3' - AATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCGACCAGTTGCAGCCCGGTTTCGAACC-5'
 * ΔR11

3' - AATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCGACCCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR21

3' - AATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR25

3' - AATGAGTTGGCATTGGCTAAAACGGTCCAATGCGCCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR26

3' - AATGAGTTGGCATTGGCTAAAACGGTCCAATCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR30

3' - AATGAGTTGGCATTGGCTAAAACGGTCCACGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR31

3' - AATGAGTTGGCATTGGCTAAAACGGTCCACGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR32

3' - AATGAGTTGGCATTGGCTAAAACGGTCCCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR33

3' - AATGAGTTGGCATTGGCTAAAACGCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR37

3' - AATGAGTTGGCATTGGCTAAAACCGGTTTCGAACCGTGACCAGGCA-5'
 * ΔR38

Figure 4.4 DNA sequence adjacent to deletion mutants ΔR11 to ΔR38. Only the DNA sequence brought adjacent to the sequence encoding the carboxy-terminal domain of the *glmS* within the 958 bp *attTn7* site is shown. Since the DNA sequence (ddTTP track only, see figure 4.2) shows no deletion within *lacZ* region of the M13 phage vector, it is presumed that the nucleotides placed next to the *attTn7* site DNA are correct. The dots indicate the positions of identical nucleotides within *attTn7* and *lacZ* DNA. DNA encoding the C-terminal domain of the *glmS* within *attTn7* site is shown in boxes. The star is the translational stop codon of *glmS*.

Therefore, a 3 fold drop in the transduction frequency in deletion mutant $\Delta R26$ in comparison with $\Delta R25$ may be associated with a variation in the measurement of transduction frequency.

By testing the function of systematic deletion mutants from the left and right end of the 958 bp attTn7 DNA, a 64 bp attTn7 region is defined which is involved in the high frequency transposition of Tn7 (see figure 4.5). The deletion mutants $\Delta R25$ and $\Delta L849$ appear to have fully functional activity with respect to attTn7, which suggests that the minimum attTn7 is an element of less than 64 bp.

These results also corroborate with transduction frequency results obtained using M13 clones containing the left and right Tn7-attTn7 site junction fragments as targets for Tn7 transposition (see chapter 3, section 3.5). It was observed in the previous chapter that the left side of the attTn7 site (with respect to the Tn7 insertion site) is less important than the right side for target site selection. The deletion analysis presented here suggests that only 3 bp of the left end of attTn7 from the point of insertion of Tn7 (5'-GCGGG-3'), in conjunction with the 56 bp of the right end, can produce the wild-type transduction frequency.

Primary <u>attTn7</u> site	phoS	Tn7 insertion site	glmS	Mean Normalised Transposition Frequencies
	—————	▼	—————	1.0E 0
ΔL124	—————		—————	8.8E-1
ΔL284	—————		—————	9.8E-1
ΔL425	—————		—————	1.0E 0
ΔL693		—————	—————	1.3E 0
ΔL843			—————	2.2E 0
ΔL849			—————	1.0E 0
ΔR11	—————		—————	1.3E 0
ΔR21	—————		—————	8.8E-1
ΔR25	—————		—————	7.5E-1
ΔR26	—————		—————	2.9E-1
ΔR30	—————		—————	8.0E-3
ΔR31	—————		—————	4.2E-3
ΔR32	—————		—————	9.8E-6
ΔR37	—————		—————	≥ 1.0E-6
ΔR45	—————		—————	≥ 1.0E-6

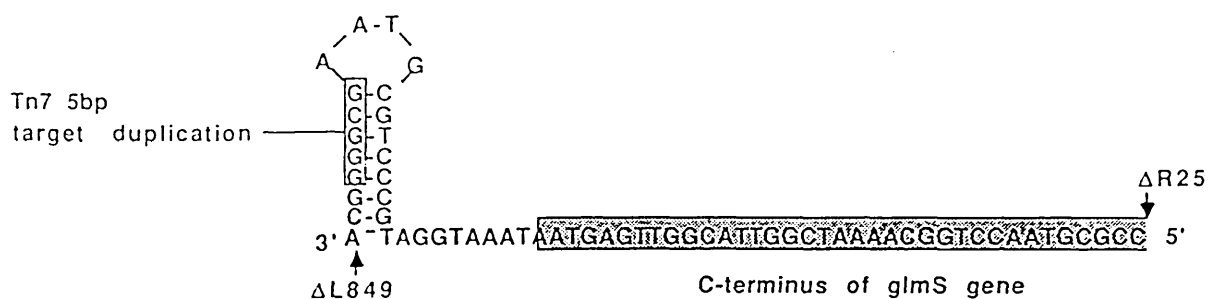


Figure 4.5 Summary of unidirectional deletion analysis of the 958 bp attTn7. The 64 bp attTn7 sequence defined by this analysis (between ΔR25 and ΔL849) is shown below which contains 36 bp sequence encoding the carboxy-terminal domain of glmS. No deletion mutants were obtained further to ΔL849. The secondary structure is shown to indicate the transcriptional terminator of the glmS gene.

The 64 bp may not be the minimum attTn7 DNA sequence required for high frequency transposition of Tn7, as deletions from the left end were not completed. The DNA sequence requirements for site-specific integration of bacteriophage lambda has been shown to include a 25 bp DNA at the attB in the chromosome of E. coli into which integrase binds to three sites (Nash, 1981).

From the data presented here, it seems that there is a requirement for a DNA sequence (not more than 64 bp) at the target site for site-specific, high frequency transposition of Tn7. However in this sequence, the 36 bp encoding the the C-terminal domain of glmS has a significant effect for determining the target site selection: a 7 bp deletion in this region abolishes the transposition of Tn7 in the transcriptional terminator of the glmS where the transposon actually inserts. It is not known if a deletion from the left end of the attTn7 site extending into the insertion site would remove the attTn7 function.

The minimal attTn7 site described above was defined by two separate deletions (Δ R25 and Δ L849, right and left end respectively). It was now important to know whether an attTn7 deleted at both ends retains full activity. The strategy for the construction of this molecule is described in section 4.4.

4.4 Cloning of a 70 bp attTn7 Site into M13 Phage Vector, mp10:

It was decided to use a HincII RE for cloning a 70 bp attTn7 DNA fragment from the deletion mutant Δ L849. A unique HincII RE site was present only 6 bp outside the 36 bp encoding the carboxy-terminal domain of glmS. This provided a favourable set of circumstances to excise a 70 bp fragment encompassing the attTn7 site for cloning purposes, which otherwise may have been difficult. Figure 4.6 represents the diagrammatic representation of the strategy adopted for cloning this 70 bp fragment (mini attTn7) in M13, mp10.

Δ L849 (RF) DNA was cleaved with HindIII and HincII REs to obtain the 70 bp fragment. The sticky HindIII ends of this fragment were filled using Klenow-polymerase and dNTP including $(\alpha\text{-}^{32}\text{P})\text{dCTP}$ as a label. The blunt ended DNA fragment so generated was isolated from a 5% polyacrylamide gel and cloned into the SmaI RE of the M13 phage vector, mp10. Colourless M13 plaques were screened by DNA sequencing (initially ddTTP reaction only) to identify the correct recombinant clones. Only 6 out of 24 recombinant phages have the correct insert and the remaining clones were unidentifiable. The complete DNA sequence of the 70 bp attTn7 site (named MAS) was determined for confirmation and figure 4.7 shows the DNA sequence of 70 bp mini attTn7 site. An autoradiograph of DNA sequencing gel of MAS is presented in figure 4.2.

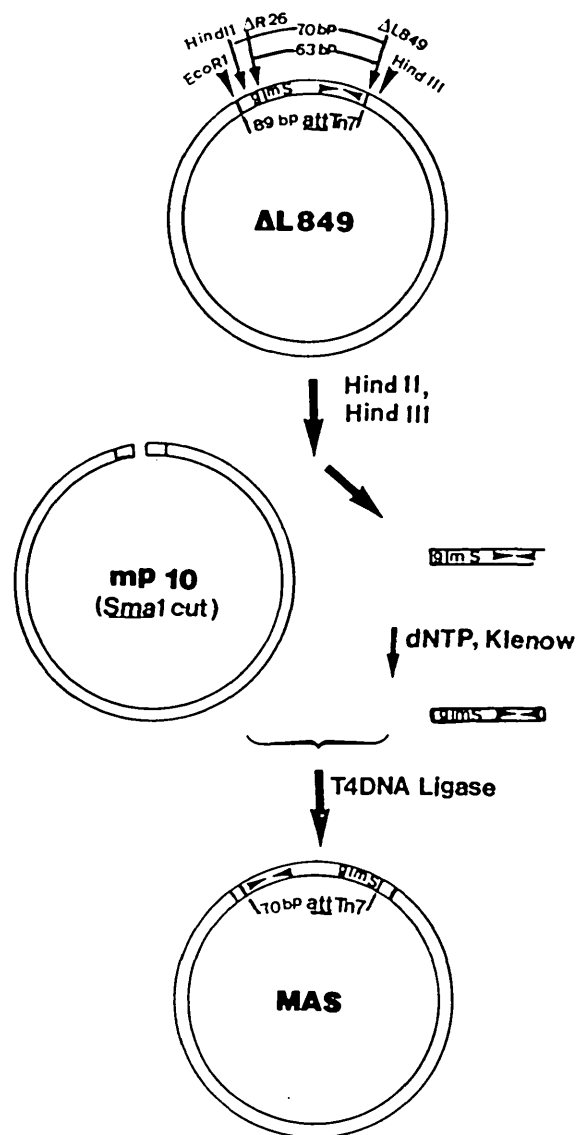


Figure 4.6 Strategy for cloning the 70 bp mini attTn7 sequence from deletion mutant Δ L849. MAS contains this fragment at the SmaI RE site of M13, mp10.

The mini attTn7 site (MAS) was tested several times for transposition of Tn7 into it by using the M13 transduction assay in a recipient bacterial strain TG2. The transduction frequency was found to be of the same order as that of VHV12 (see table 4.2).

Table 4.2 Results of the M13 transduction frequency assays on 70 bp mini. attTn7 site (MAS).

Independent trials	Transduction frequency		Relative to VHV12
	MAS	VHV12	
1	1.3E-4	1.1E-4	1.2
2	6.4E-5	7.0E-5	0.9
3	1.0E-4	1.2E-5	0.8
4	2.5E-4	1.0E-4	2.5

The transduction frequency results were consistent with the hypothesis that the minimum attTn7 DNA information is contained within the 70 bp attTn7 fragment and this fragment can serve, independent of the 958 bp attTn7, as an efficient target site for Tn7 transposition.

However, in this sequence only the right end of the attTn7 site was defined as in the left hand end deletions beyond position 849 were not obtained. Therefore, it is premature to assess the involvement of the sequence further from the left end of the 70 bp attTn7. The potential involvement of any specific DNA sequence, within the 70 bp attTn7 site for regulating the high frequency transposition of Tn7, is examined in the remaining chapters.

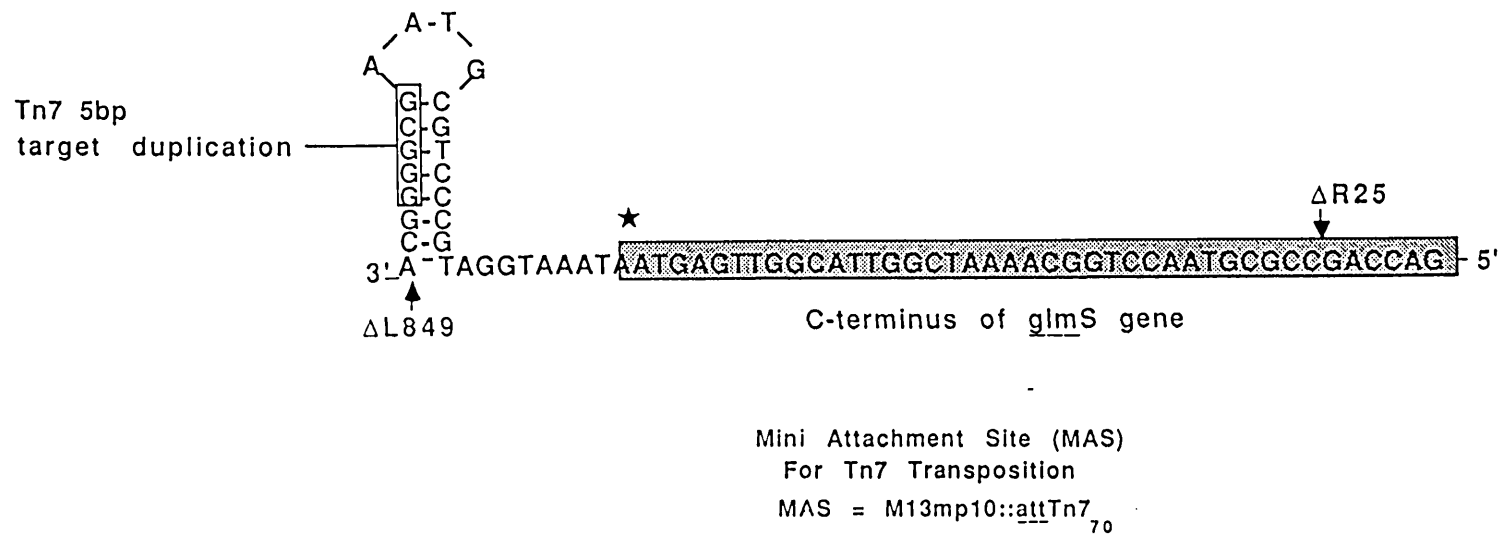


Figure 4.7 DNA sequence of the 70 bp cloned mini attTn7 site. In this region 36 bp of the sequence encoding carboxy-terminal domain of glmS seems to be important in regulating the high frequency transposition of Tn7 in the glmS transcriptional terminator. This sequence is between the arrow at position ΔR25 and the star at position 61.

CHAPTER 5

Analysis of Tn7 Transposition into the Chromosome of Agrobacterium tumefaciens, Klebsiella pneumoniae, Serratia marcescens and to Secondary Sites in Escherichia coli

5.1 Introduction:

High efficiency and specificity of Tn7 transposition has been shown to occur in the chromosomes of several bacteria in addition to E. coli (see chapter 1 section 1.10). However, the DNA sequence and chromosomal location of attTn7 has only been characterised in E. coli (Lichtenstein and Brenner, 1982).

In this study, three other bacterial species were studied which had not previously been shown to have high frequency, high specificity Tn7 insertion sites (attTn7 sites). They were chosen on the basis of varying degrees of relatedness to E. coli. One species selected K. pneumoniae is a very close relative of E. coli. Another, S. marcescens, though more distantly related, is within the same family, the Enterobacteriaceae. The third species, A. tumefaciens, is in another (distantly related) family, the Rhizobiaceae (see figure 5.1).

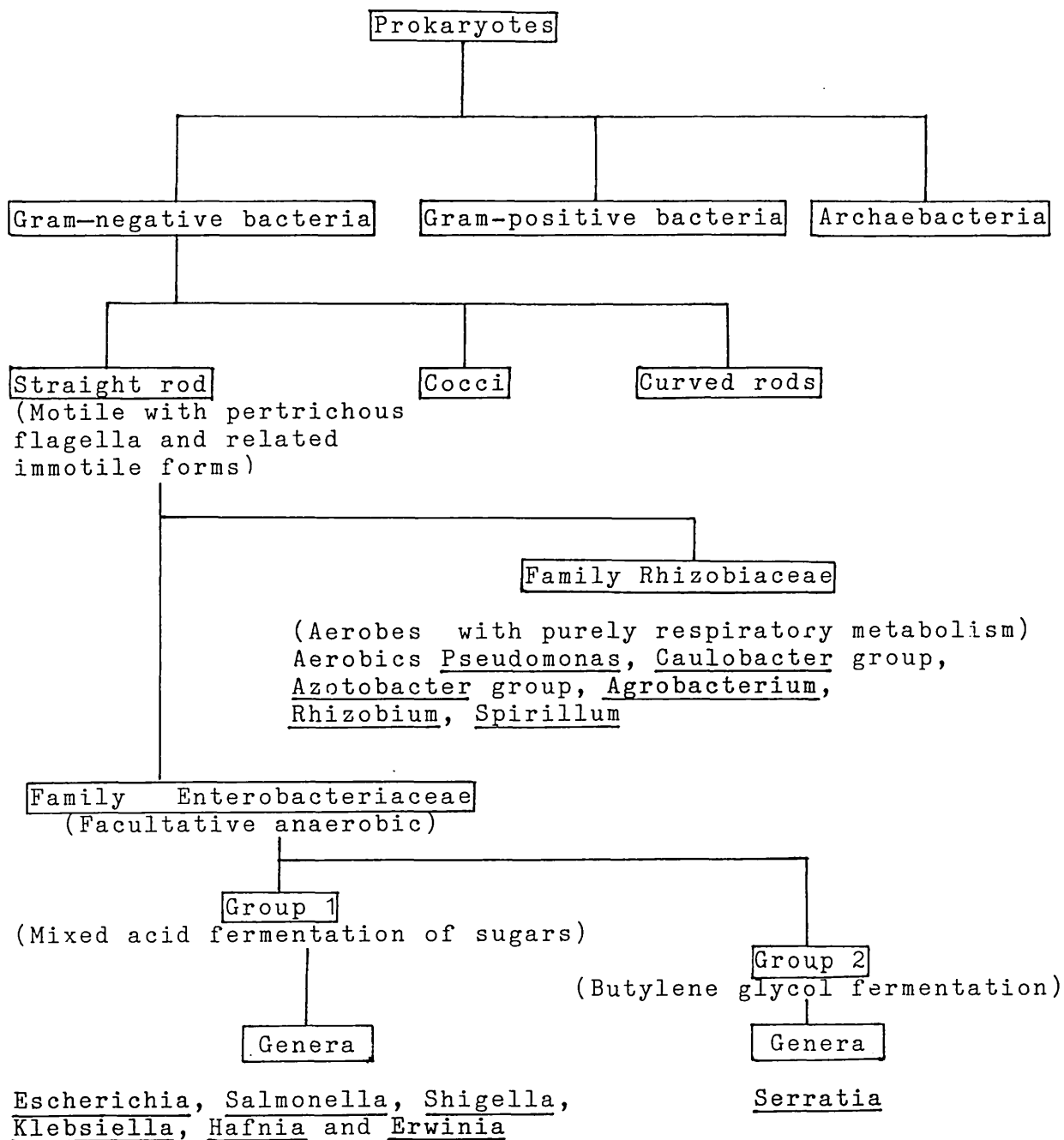


Figure 5.1 Phlyogenetic relationship between the bacterial species studied for site-specific transposition of Tn7.

In choosing these species, four questions were addressed: First, what is the frequency of Tn7 transposition into the bacterial chromosome? Second, does Tn7 also transpose to unique insertion sites? Third, what is the DNA sequence of these sites? Fourth, are there secondary attachment sites for Tn7 transposition in the E. coli chromosome? The third question is addressed in the following chapter.

5.2 Experimental Strategy for the Selection for Tn7 Transposition to the Bacterial Chromosome:

In order to develop a selection strategy for Tn7 transposition to the bacterial chromosome, it was clearly necessary to know whether Tn7 encoded drug resistance markers (Tp^r , Sm/Sp^r) could be used as the basis for such a selection. Thus the first requirement was information on the drug sensitivity of these bacteria. A. tumefaciens, K. pneumoniae and S. marcescens strains were tested on varying concentrations of trimethoprim and spectinomycin to determine the minimum inhibitory concentration at which the growth of a single colony was completely inhibited.

All, except A. tumefaciens, were found to be Tp^s . The mechanism of resistance of A. tumefaciens to trimethoprim was not established. It may be that the antibiotic is not transported into the cell. It is also possible that trimethoprim does not inhibit DHFR activity of A. tumefaciens. If so, the A. tumefaciens gene encoding DHFR

cannot be identical to that encoded by Tn7 as evidenced by DNA hybridisation studies described later in this chapter. A third possibility is that an endogenous drug resistance gene exists which inactivates this antibiotic. As trimethoprim provides a stronger selection it was used for K. pneumoniae and S. marcescens but spectinomycin was used for A. tumefaciens in all subsequent experiments.

The second requirement was a method to deliver Tn7 to these bacteria on a plasmid donor which could be removed; selecting for loss of this plasmid and asking for retention of Tn7 should only result from Tn7 transposition to the chromosome, providing Tn7 encoded drug resistance genes are expressed in these bacteria.

Two approaches were taken to meet this requirement. The first involved the use of a suicide Tn7 donor vector unable to replicate in these bacteria; plasmids with ColE1 type origins of DNA replication (such as pBR322) have very narrow host ranges and, for example, cannot replicate in A. tumefaciens (Lichtenstein and Draper, 1984). It was not known whether or not such plasmids replicate in K. pneumoniae and S. marcescens. In order to ask whether the pBR322 derived plasmid pCPL6::Tn7 (Lichtenstein and Brenner, 1982) could be used as a suicide donor of Tn7, **CCC** pCPL6::Tn7 plasmid DNA was used for transformation into A. tumefaciens, K. pneumoniae and S. marcescens using the freeze/thaw protocol to prepare competent cells (Lichtenstein and Draper, 1984).

However, no Tp^r colonies (in S. marcescens and K. pneumoniae) and Sp^r colonies (in A. tumefaciens) were obtained. This could be due to several reasons; (1) low frequency transformation, (2) plasmids may have entered cells but perhaps were unable to replicate and were degraded before Tn7 transposition could take place, and (3) Tn7 encoded genes are not expressed in these strains.

This approach was abandoned in favour of the second approach exploiting plasmid incompatibility to chase out the Tn7 donor plasmid with a second incompatible plasmid. Here plasmids were introduced to target bacterial strains by conjugation in two steps as described by Miller (1972). The diagrammatic representation of this strategy is shown in figure 5.2.

5.2.1 Step-1; Conjugal Transfer of RP4::Tn7:

The large (49 Kb) wide host range conjugative plasmid RP4::Tn7 (encoding Kn^r , Tc^r , Ap^r in addition to Tn7 encoded markers) was chosen as the Tn7 donor. The Pseudomonas aeruginosa strain, PAS279, was used as the donor of RP4::Tn7 and conjugated with the three recipients. Conjugations were performed on non-selective YT plates by tooth-picking donor and recipients together.

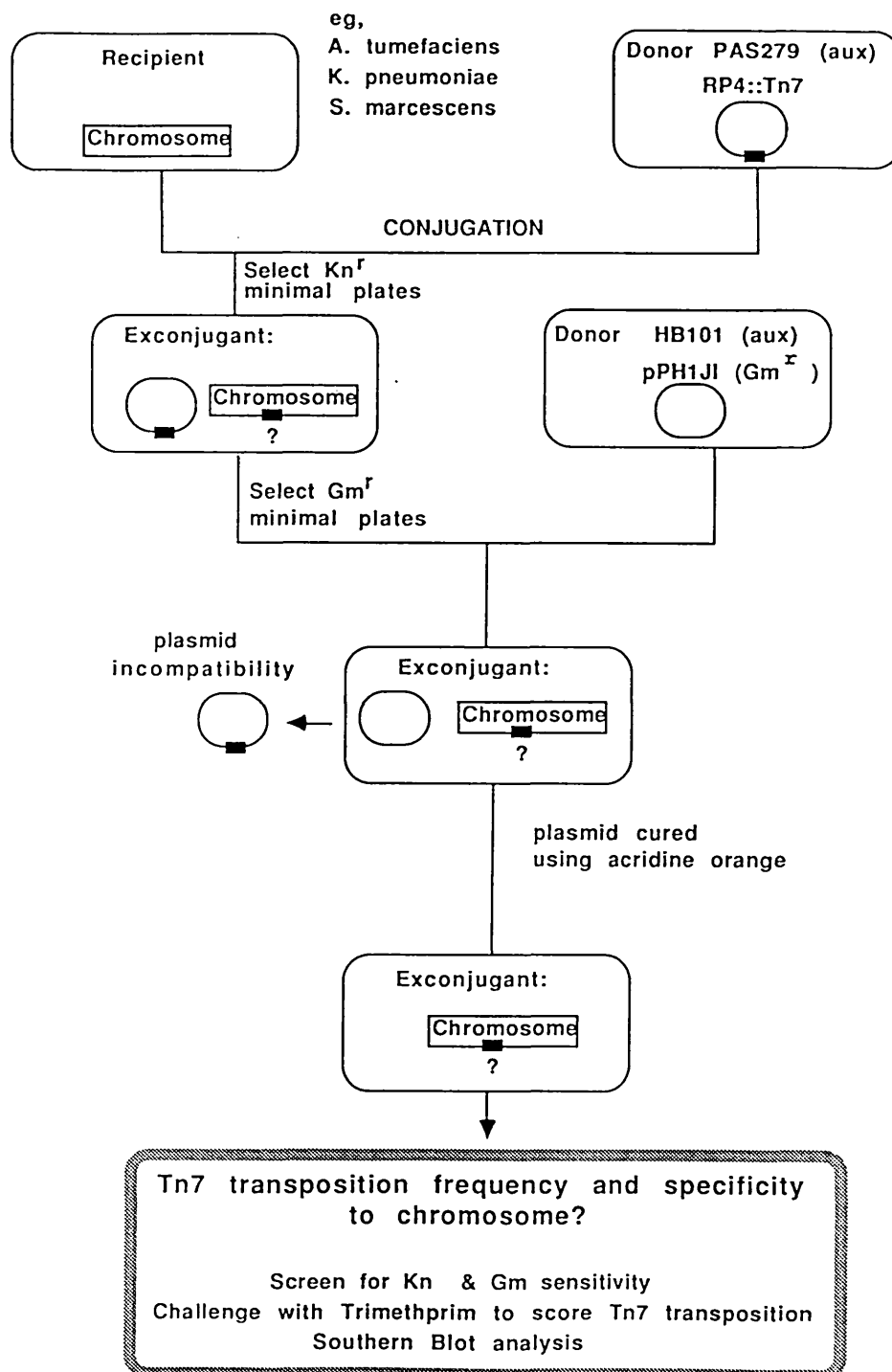


Figure 5.2 Diagrammatic representation of experimental strategy for the selection for Tn7 transposition to the bacterial chromosome.

Four hundred independent matings of each strain were performed in order to obtain a large number of independent Tn7 transposition events. Then exconjugants were selected by tooth-picking the 400 independent mixed colony matings onto minimal AB plates containing 100µg/ml kanamycin. The recipients had previously been tested for prototrophy and ^sKn . The donor PAS279 is auxotrophic requiring threonine for growth and thus cannot grow on AB plates.

5.2.2 Step-2; Exclusion of RP4::Tn7 by Conjugal Transfer of Incompatible Plasmid pPH1JI:

The mobilisation of pPH1JI was designed to exclude the plasmid RP4::Tn7 from the three bacterial strains containing RP4::Tn7. Both plasmids belong to the incP group. Like plasmids of incC, incN, incQ and inc W, incP plasmids have broad host ranges and replicate stably in a wide variety of Gram negative bacteria. However, two plasmids of the same incompatibility group cannot coexist in the same bacterial cell since their replication systems are compatible. The incP plasmid pPH1JI (54.7 Kb) was constructed by Hirsch (1978) for such plasmid chase experiments. This plasmid, which encodes gentamycin resistance (^rGm) but not ^rKn , ^rTc or ^rAp was constructed to exclude many other incP plasmids.

The three sets of bacterial exconjugant strains which contain the RP4::Tn7^r were selected for Gm^r following the mating with the E. coli donor HB101 (pPH1JI). This donor is an auxotrophic strain needing histidine and proline for growth on minimal plates and thus provided an easy selection for pPH1JI delivery to the three sets of recipients. Subsequent to another tooth-pick mating four hundred independent exconjugant colonies were selected on minimal plates containing 100µg/ml gentamycin. Bacterial colonies were picked successively onto such plates at least 10 times to select for replacement of RP4::Tn7 by pPH1JI.

The exconjugants were then tested for kanamycin sensitivity screening for loss of RP4::Tn7. Thus gentamycin resistant exconjugants sub-cultured in this way were replica tooth-picked onto AB Kn and AB Tp (K. pneumoniae, S. marcescens) or AB Sp (A. tumefaciens) plates. Kanamycin sensitive colonies able to grow on the AB Tp or AB Sp plates suggest that a Tn7 transposition event from RP4::Tn7 has occurred. This frequency of transposition of Tn7 is therefore expressed as the ratio of;

$$\left[\frac{\text{Number of Tp}^{\text{r}} \text{ (or sp)}^{\text{r}}, \text{Kn}^{\text{s}} \text{ colonies.}}{\text{Total number of Kn colonies.}^{\text{s}}} \right]$$

Of the original 400 exconjugants for each species (which had acquired RP4::Tn7 in step-1) more than 95% had lost kanamycin resistance subsequent to the plasmid chase using pPH1JI (step-2), yet almost all of them retained Tn7

encoded markers. These positive results suggest that Tn7 transposes with high frequency from the donor plasmid RP4::Tn7 to new target sites. At this stage the target for Tn7 transposition was not established.

5.3 Measurement of Tn7 Transposition Frequency:

In order to obtain the exact ratio of the event described above it was important to analyse independent streak purified colonies at each stage of the mating rather than using the mixed mating colonies. Therefore 10 colonies of each species was taken at a time and mated with the donor RP4::Tn7. The colonies were separated from the mating mixture by growing them onto AB plates containing Kn. The 10 individually separated exconjugants were mated with the HB101 (pPH1JI) and the exconjugants were then again streaked purified on AB plates containing Gm. Individual colonies from each species were repeatedly selected for the loss of RP4::Tn7 (Kn^s) and retention of the pPH1JI (Gm^r). The cumulative results of these matings for each bacterial strain are presented below.

	$\text{Transposition frequency} = \frac{\text{Number of Tp or Sp, Kn colonies}^{\text{r}}}{\text{Number of Kn colonies}^{\text{s}}}$	
<u>A. tumefaciens</u>	48/50	(96%)
<u>K. pneumoniae</u>	38/40	(95%)
<u>S. marcescens</u>	45/50	(90%)

The most important feature of this assay was the natural selection of Tn7 transposition from the Tn7 donor replicon to new sites in the three bacterial species, independent of selection for Tn7 encoded antibiotics. Therefore, there was no bias in establishing the Tn7 transposition frequency results. In all cases the transposition frequency of Tn7 was very high, similar to the observed values of Tn7 transposition frequencies in other bacterial species (Barth et al., 1976; Datta et al., 1980; Lichtenstein and Brenner, 1981).

5.4 Determination of Tn7 Target Replicon:

The series of events which could take place at the end of stage-2 are the transposition of Tn7 either into the chromosome of these bacterial species or into the plasmid pPH1JI. The study of the nature of attTn7 requires the construction of strains, each containing the independent insertion of Tn7 in the chromosome. Therefore, it was important to cure the plasmid pPH1JI from these strains to establish that Tn7 had transposed to the chromosome. The curing was performed by using varying concentrations of acridine orange on these strains. The procedure for this treatment is described in chapter 2, section 2.2.1(v). Acridine orange interchelates with the CCC supercoiled DNA (plasmid or episome) with high affinity in comparison to the chromosomal DNA and thus blocks the replication of plasmids or episomes.

Forty independent colonies of each species were repeatedly grown (at least 10 times) by tooth-picking a single colony from the AB plates containing acridine orange to new plates and incubated at the appropriate temperature. After this selection, each colony was then replica tooth-picked onto AB, Gm plates and AB, Tp (K. pneumoniae and S. marcescens) or AB, Sp (A. tumefaciens) plates. All colonies were unable to grow on AB Gm plates and therefore had lost the plasmid pPH1JI. The same colonies were able to grow successfully on AB, Tp or AB Sp plates, indicating the retention of Tn7 and arguing that Tn7 transposition was to the chromosome. Since in the 40 colonies the loss of plasmid pPH1JI was not related to the transposition of Tn7 in this plasmid, it is likely be the same result in the remaining 360 colonies. These results have clearly suggested the high frequency transposition of Tn7 into the chromosome of these bacteria from the Tn7 donor, replicon RP4::Tn7.

During the course of this work, a summer student working under my guidance also found high frequency transposition of Tn7 in the chromosome of Erwinia caratovora, a member of the family Enterobacteriaceae.

5.5 Southern Blot Analysis to Determine Site-specificity of Tn7 Transposition to the Chromosome:

Does Tn7 transpose to unique sites in these bacterial chromosomes? This question was approached biochemically using the method of Southern (1975) to blot restriction fragments of total chromosomal DNAs which were isolated from strains derived from independent Tn7 transposition events. The blots were probed with radioactively labelled Tn7 DNAs to identify those restriction fragments which include Tn7 DNA.

To determine the orientation of inserted Tn7 in the chromosome, plasmid pRS5 DNA (which has a deletion of the large, 8.9 Kb EcoR1 RE fragment from ColE1::Tn7) or the purified right end (8.9 Kb) EcoR1 RE fragment of Tn7, were used as hybridisation probes. Thus, total chromosomal DNA digested with EcoR1 should give two restriction fragments with radiolabelled ColE1::Tn7 probe and only one band of homology while using the radiolabelled pRS5 or right end (8.9 Kb) fragment of Tn7.

Genomic DNA was isolated from 9 independent colonies of A. tumefaciens and S. marcescens and also 10 independent colonies of K. pneumoniae with independent insertions of Tn7 into the chromosome and a control without insertion of Tn7. Plasmid pTJS140 was used as a control to show that plasmid ColE1 of the hybridisation probe did not have homology with the RP4 derivative plasmid, as it is a hybrid plasmid containing the ColE1 and RP4 DNA. The

HindIII restricted pTJS140 yields two fragments; 3.2 Kb of ColE1 origin and 4.5 Kb of RP4 origin. It was expected that plasmid ColE1 would not show homology with the 4.5 Kb HindIII RE fragment of pTJS140, thus in the hybridisation probe (ColE1::Tn7), only Tn7 homomlogy will appear in genomic DNA (chromosomal DNA::Tn7) on the autoradiograph.

4-5µg of genomic DNA of each isolate of each bacterial species was digested with EcoRI RE and loaded on a 0.8% agarose gel, along with 0.5µg of plasmid pTJS140 digested with HindIII RE and 0.2µg of ColE1::Tn7 DNA digested with EcoRI as a positive control. The fractionated digests were blotted onto nitrocellulose filters. The loading and transfer of DNA was done in duplicate for probing with ColE1::Tn7 and pRS5 and the filters hybridised with each probe were autoradiographed. The results of these blots are shown in figures 5.3, 5.4 (for A. tumefaciens), 5.5, 5.6 (for K. pneumoniae) and 5.7, 5.8 (for S. marcescens).

The results obtained for A. tumefaciens and K. pneumoniae are clear: in both strains two bands [of estimated size (19 Kb and 5.3 Kb) and (16 Kb and 9.5 Kb) respectively ($\pm$ 0.5 Kb)] of homology appeared for the set of 9 and 10 strains using radiolabelled ColE1::Tn7 as a probe. In A. tumefaciens and K. pneumoniae the probe pRS5 hybridised with one band (i.e. 5.3 Kb and 9.5 Kb respectively) but not with the second band (i.e. 19 Kb and 16 Kb respectively). Therefore it is clear that insertion of Tn7

in the chromosome of A. tumefaciens and K. pneumoniae occurs at a specific site and in one orientation in all independent Tn7 transpositions events. A similar observation of high frequency and site-specific transposition of Tn7 with unique orientation in the chromosome of Erwinia caratovora was made (data not shown).

However, the results in S. marcescens were less clear. In all nine independent events there seems to be four bands [of estimated sizes 19Kb, 10 Kb, 7 Kb and 5.3 Kb ($\pm$ 0.5 kb)] of homology with radiolabelled ColE1::Tn7 and all these four bands again have homology with the right end (8.9 Kb) fragment of Tn7 (figures 5.6 and 5.7). The 4 EcoRI RE fragments of the genomic DNA had the same degree of homology while using the radiolabelled pRS5 DNA as a hybridisation probe (data not shown). The filters were washed in high stringency conditions [described in chapter 2, section 2.2.2F(i)], therefore, the possibility of homology to the both ends of Tn7 in the EcoRI RE digested genomic DNA was very small. This is clear in the control ColE1::Tn7 that only one out of two EcoRI fragments shows homology with the probe (see figure 5.8). Interestingly, in the event number 5, the third band (7 Kb) did not show any homology with the left end probe. One reason may be a block in transferring the DNA at this particular position on the nitrocellulose filter. Other possible reasons are discussed later in section 6.

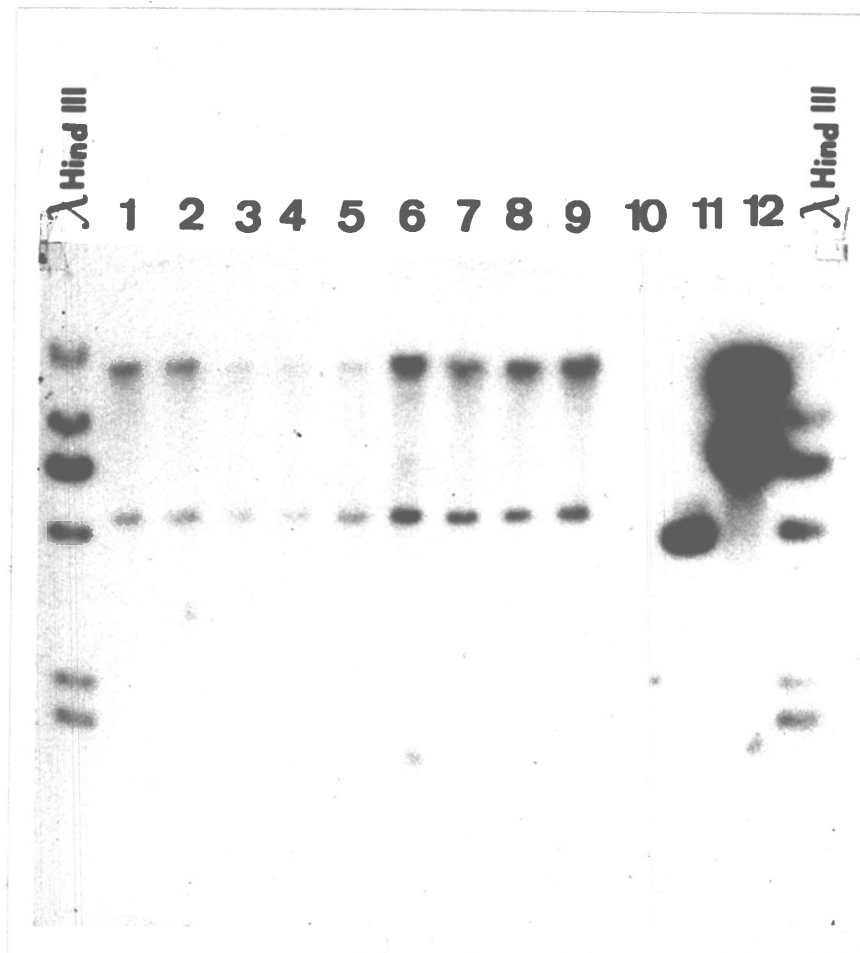


Figure 5.3 Southern Blot of EcoRI RE digested genomic DNA of A. tumefaciens::Tn7 probed with radiolabelled ColE1::Tn7 DNA. This blot shows 9 independent insertions of Tn7, A. tumefaciens::Tn7 chromosomal DNA EcoRI RE digests (lanes 1-9); A. tumefaciens DNA, EcoRI RE digest (lane 10); pTJS140, HindIII RE digest (lane 11) and ColE1::Tn7 EcoRI digest (lane 12). In all cases it is clear that Tn7 has inserted at a unique site in the chromosome, thus showing the two EcoRI RE fragments of the identical sizes containing the Tn7-attTn7 DNA (of estimated size 19 Kb upper and 5.3 Kb lower fragment).

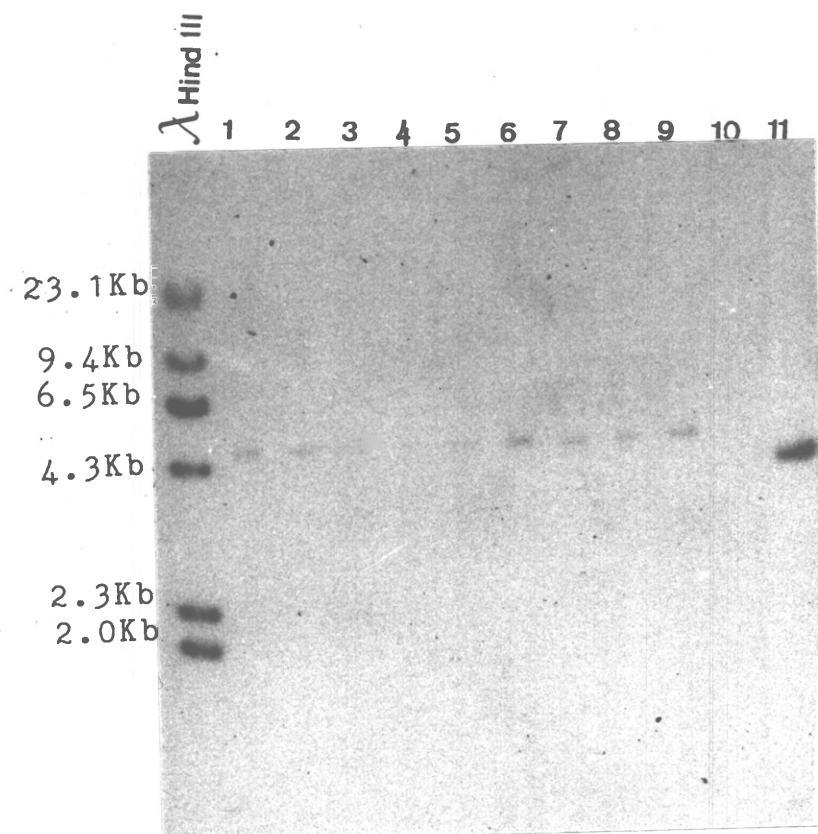


Figure 5.4 Southern Blot of EcoRI RE digested genomic DNA of A. tumefaciens::Tn7 probed with radiolabelled pRS5 DNA. The loading pattern is identical to that of figure 5.3. In all cases the plasmid pRS5 shows homology with the 19 Kb EcoRI RE fragment of Tn7-attTn7 but not with the 5.3 Kb fragment, thus suggesting that Tn7 insertion has occurred in one orientation.

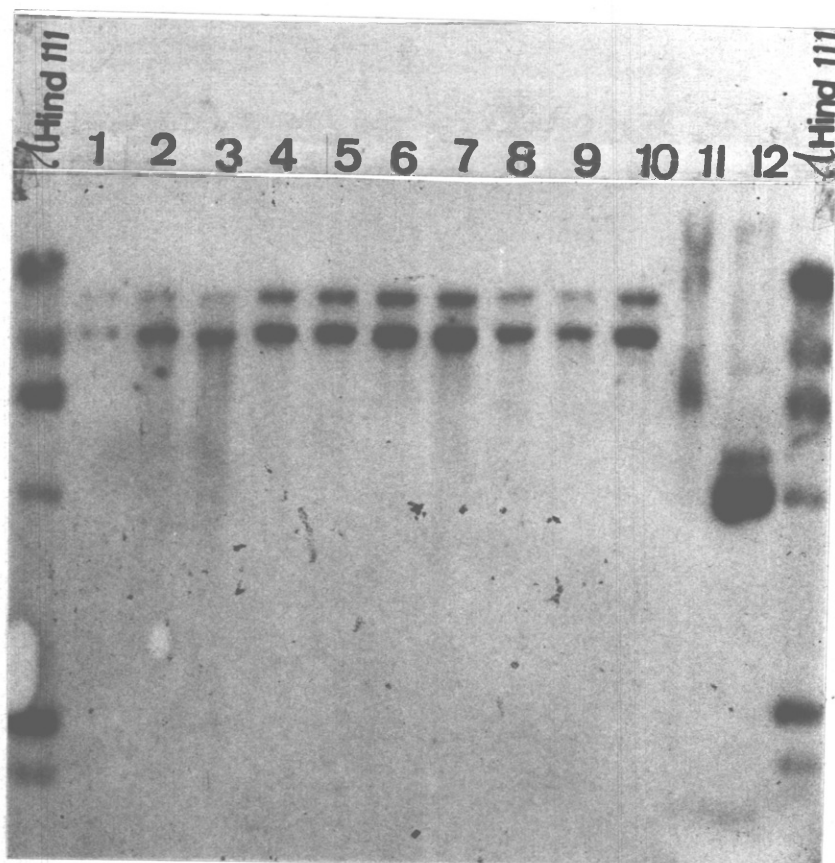


Figure 5.5 Southern Blot of EcoRI RE digested genomic DNA of K. pneumoniae::Tn7 probed with radiolabelled ColE1::Tn7 DNA. This blot shows 10 independent insertions of Tn7, K. pneumoniae::Tn7 chromosomal DNA EcoRI RE digests (lanes 1-10); K. pneumoniae DNA, EcoRI RE digest (lane 11); pTJS140, HindIII RE digest (lane 12). In all cases it is clear that Tn7 has inserted at a unique site in the chromosome, thus showing the two EcoRI RE fragments of the identical sizes containing the Tn7-attTn7 DNA (of estimated size 16 Kb upper and 9.5 Kb lower fragment).



Figure 5.6 Southern Blot of EcoRI RE digested genomic DNA of K. pneumoniae::Tn7 probed with radiolabelled pRS5 DNA. The loading pattern is identical to that of figure 5.5. In all cases the plasmid pRS5 shows homology with the 9.5 Kb EcoRI RE fragment of Tn7-attTn7 but not with the 16 Kb fragment, thus suggesting that Tn7 insertion has occurred in one orientation.

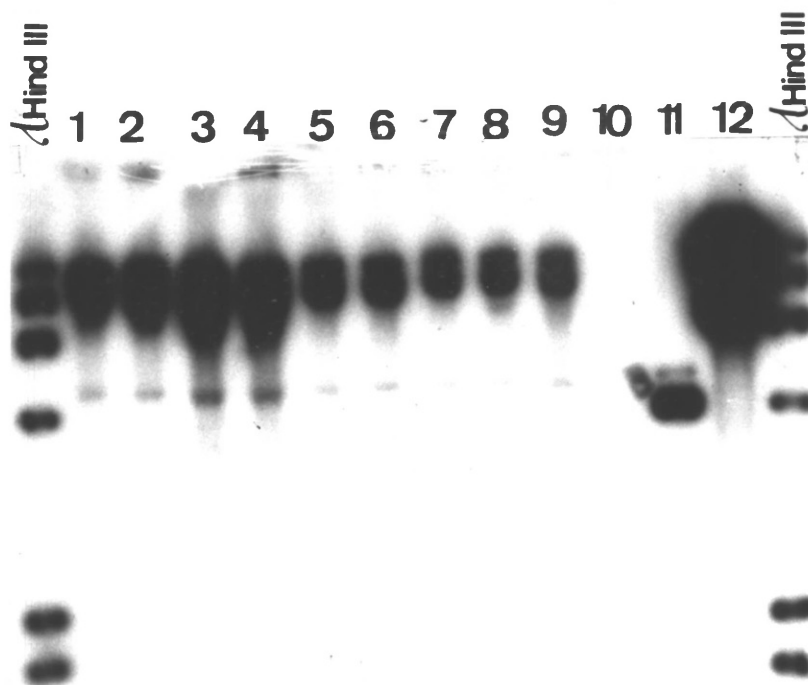


Figure 5.7 Southern Blot of EcoRI RE digested genomic DNA of S. marcescens::Tn7 probed with radiolabelled Cole1::Tn7 DNA. This blot shows 9 independent insertions of Tn7, S. marcescens Tn7 chromosomal DNA EcoRI RE digests (lanes 1-9); S.marcescens DNA, EcoRI RE digest (lane 10); pTJS140, HindIII RE digest (lane 11) and Cole1::Tn7 EcoRI digest (lane 12). The filter was exposed for 12 hours in autoradiography to define the lower band (5.1 Kb) of homology. Consequently, the upper three bands are too overexposed to differentiate in this autoradiograph, but can be seen in figure 5.8.

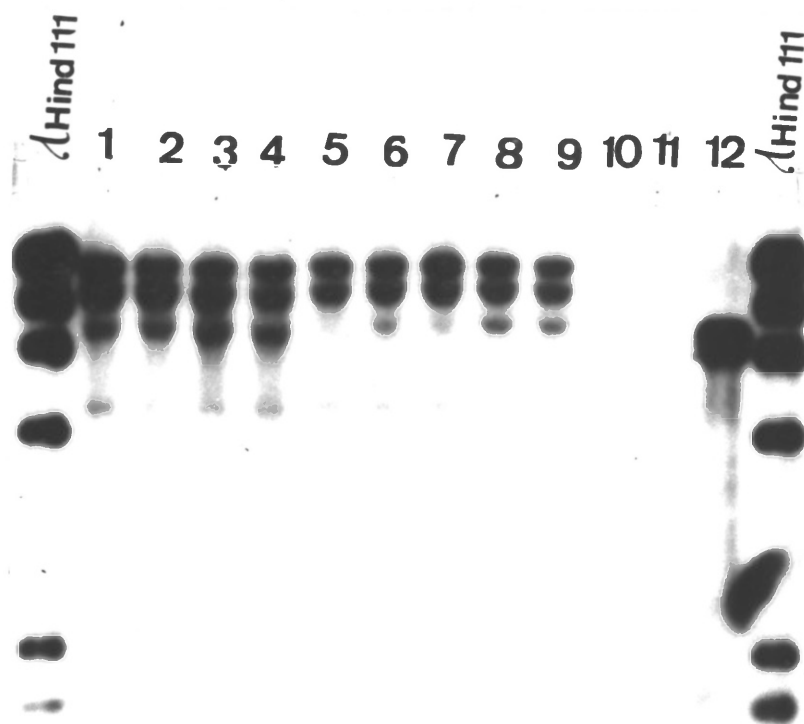


Figure 5.8 Southern blot of EcoRI RE digested genomic DNA of S. marcescens::Tn7 probed with radiolabelled EcoRI left end fragment (8.9 Kb) of Tn7 DNA. In this autoradiograph the left end of Tn7 (8.9 Kb, EcoRI fragment) shows homology with same four bands as seen in figure 5.7. The event number 5 shows only three fragments of homology. The radiolabelled probe (right end of Tn7) displays homology with only one out of two EcoRI fragments of ColE1::Tn7 suggesting that washing conditions were very stringent.

In order to determine the integrity of Tn7 in the chromosome of these species, further Southern blot analyses were carried out using the HindIII and BamH1 (restriction enzymes which release internal Tn7 fragments) cleaved chromosomal DNA of one event of insertion of Tn7 into the chromosome.

HindIII digestion of Tn7 generates two internal restriction fragments and the results shown in figure 5.9 provided the evidence that in all cases these two fragments were preserved. It can also be seen in this figure that there are four flanking fragments of estimated sizes 21 Kb, 16 Kb, 10 Kb and 6.5 Kb ($\pm$ 0.5 Kb) in the HindIII digests of S. marcescens and only two in the cases of A. tumefaciens and K. pneumoniae. This result provided further evidence that there is one insertion of Tn7 in the chromosomes of A. tumefaciens and K. pneumoniae, and two in the chromosome of S. marcescens.

The next objective of this project was to clone the attTn7 sites from the chromosomes of these species. As Tn7 encodes selectable markers a direct approach to sub-cloning Tn7 and flanking chromosomal DNA was selected (see chapter 6). This strategy requires the use of an RE which has no sites in Tn7. Therefore, to find a suitable fragment size for cloning purposes, a Southern blot was performed on the chromosomal DNA cleaved with restriction enzymes which have no sites within Tn7.

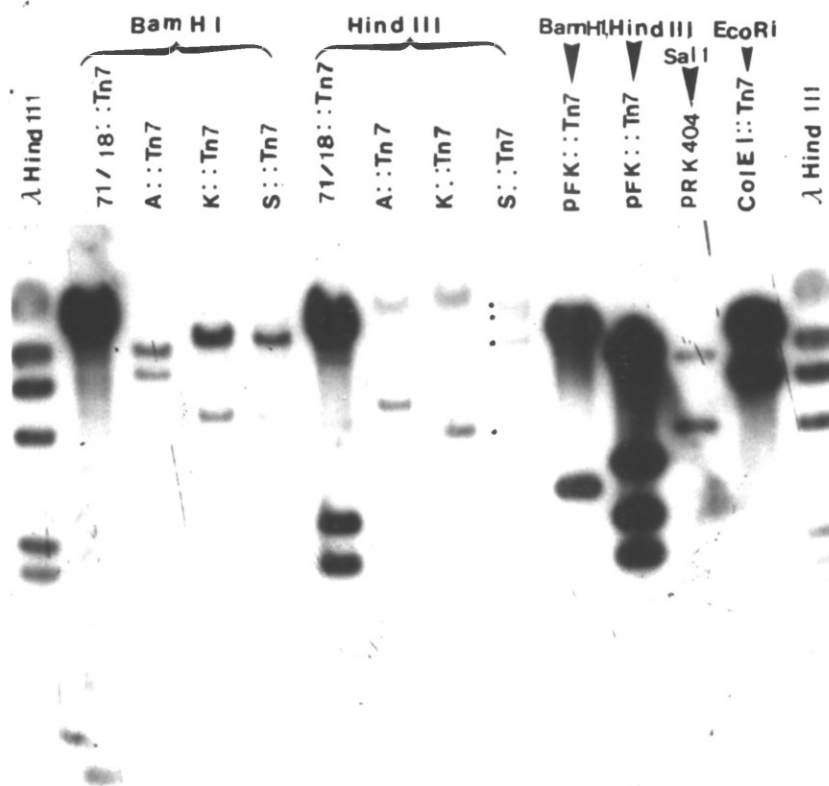


Figure 5.9 Southern blot of *Hind*III, *Bam*HI REs digested genomic DNAs of *A. tumefaciens*::Tn7 (A::Tn7), *K. pneumoniae*::Tn7 (K::Tn7) and *S. marcescens*::Tn7(S::Tn7) probed with radiolabelled *ColE1*::Tn7 DNA. The controls in this experiment are an *E. coli*::Tn7 strain (71/18::Tn7) and plasmid pFK::Tn7 digested with the same REs. The *Hind*III RE digested *S. marcescens*::Tn7 DNA shows four junction fragments of Tn7-*att*Tn7 DNA of estimated sizes 24, 22, 10 (± 0.5 Kb) and 3.5 Kb (± 0.2 Kb) as shown by the dots.

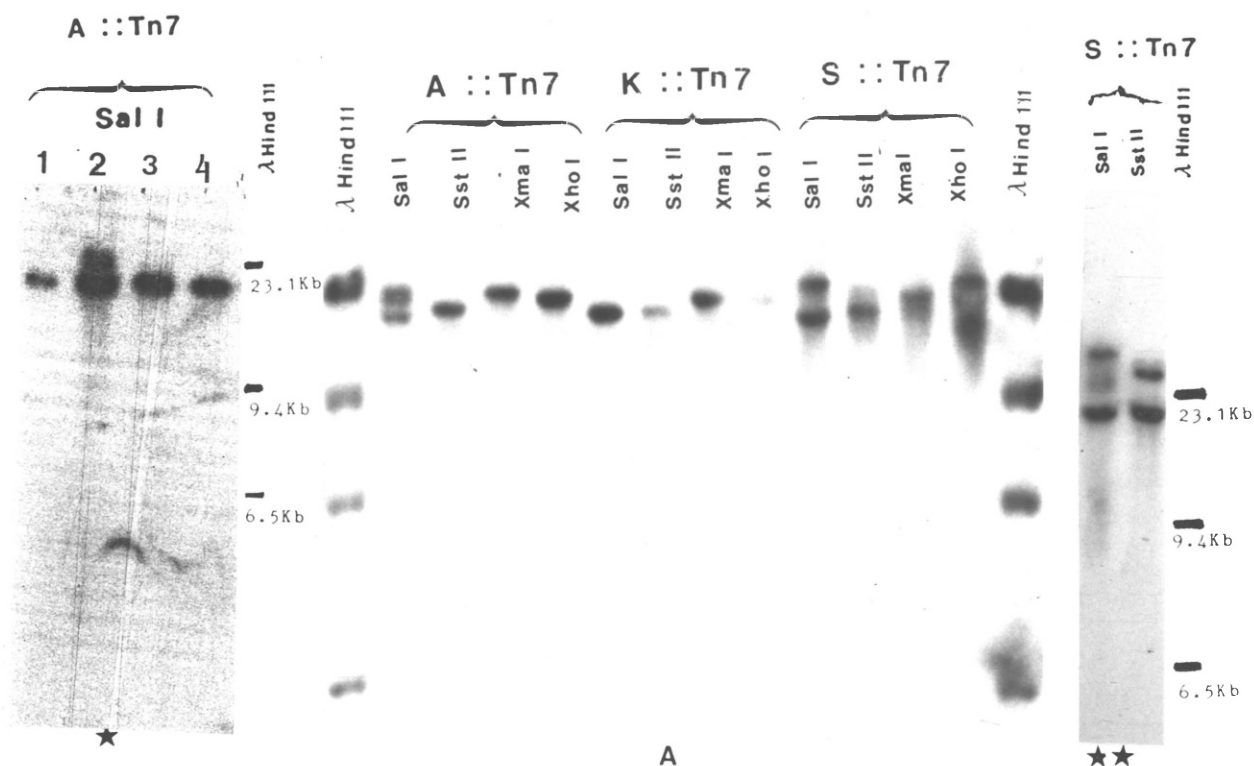


Figure 5.10 Southern blot of SalI, SstII, SmaI and XhoI REs digested genomic DNAs of A. tumefaciens::Tn7 (A::Tn7), K. pneumoniae::Tn7 (K::Tn7) and S. marcescens::Tn7 (S::Tn7); probed with radiolabelled ColE1::Tn7 DNA.

★Four independent events of Tn7 insertion in A. tumefaciens, digested with SalI and probed with radiolabelled ColE1::Tn7 DNA. This is to show that in figure (A), A. tumefaciens::Tn7 genomic DNA was partially cleaved with SalI RE.

★★One independent insertion of Tn7 in the chromosome of S. marcescens cleaved with REs SalI and SstII for hybridising with radiolabelled ColE1::Tn7. This autoradiograph is to show that Tn7 has inserted at two places in S. marcescens genomic DNA.

Gosti-Testu et al. (1983) have shown that Tn7 has no restriction endonuclease sites for SalI, SstII, SmaI/XmaI or XhoI. This observation was used to find suitable restriction fragments of Tn7 and Tn7 flanking DNA from one event of Tn7 insertion in the chromosome of each species which could be exploited later for cloning purposes. To do this, genomic DNA was cleaved with SalI, SstII, SmaI and XhoI REs and probed with radiolabelled ColE1::Tn7. The results of this blot are presented in figure 5.10. In A. tumefaciens one band of estimated size 21 Kb, ($\pm$ 0.5 Kb) for SstII, 24 Kb, ($\pm$ 0.5 Kb) for SmaI and 22 Kb, ($\pm$ 0.5 Kb) for XhoI RE digestion were observed. However, the two bands of estimated sizes 23 Kb, ($\pm$ 0.5 Kb) and 20 Kb, ($\pm$ 0.5 Kb) seen for SalI RE digestion were due to the partial digestion of the genomic DNA. This was later shown by a Southern blot on the chromosomal DNA::Tn7 of four independent events cleaved with SalI RE and in all cases one fragment of estimated size 20 Kb ($\pm$ 0.5 Kb) was shown to have homology with the radiolabelled ColE1::Tn7.

In K. pneumoniae one band of homology each of estimated size of 21 Kb, ($\pm$ 0.5 Kb) for SalI, 22 Kb, ($\pm$ 0.5 Kb) for SstII, 22.5 Kb, ($\pm$ 0.5 Kb) for SmaI and 23 Kb, ($\pm$ 0.5 Kb) for XhoI, had homology with the radiolabelled ColE1::Tn7 DNA. These results indicate that Tn7 had inserted at one site in the chromosome of A. tumefaciens and K. pneumoniae.

In S. marcescens, a different restriction pattern was seen in the genomic DNA with insertion of Tn7 when digested

with SalI, SstII and XhoI REs and probed with radiolabelled ColE1::Tn7 DNA (see figure 5.10). The SalI, SstII and XhoI REs have produced two fragments of estimated sizes 25 Kb and 20 Kb, ($\pm$ 0.5 Kb), 25 Kb and 20 Kb, ($\pm$ 0.5 Kb), 24 Kb and 16 Kb, ($\pm$ 0.5Kb) respectively (figure 5.10). The Southern blot was repeated with different events of Tn7 insertion in S. marcescens and the same results were obtained.

5.6 Tn7 has Two Insertion Sites (attTn7 sites) in S. marcescens:

From the previous section, it seemed that Tn7 has two attTn7 sites in the S. marcescens genome. The orientation of Tn7 in either case could not be determined as the hybridisation probe at either end of Tn7 showed a cross homology (i.e. the left end of Tn7 showed homology with the right end of Tn7 and vice versa). Therefore for a better understanding of Tn7 insertional properties in the chromosome of S. marcescens, 50 independent bacterial matings (as described in section 5.2.1 and 5.2.2) were performed again. Chromosomal DNA of strains with Tn7 insertions in the genome of S. marcescens was isolated and cleaved with SstII, HindIII and SalI REs for Southern blot analysis.

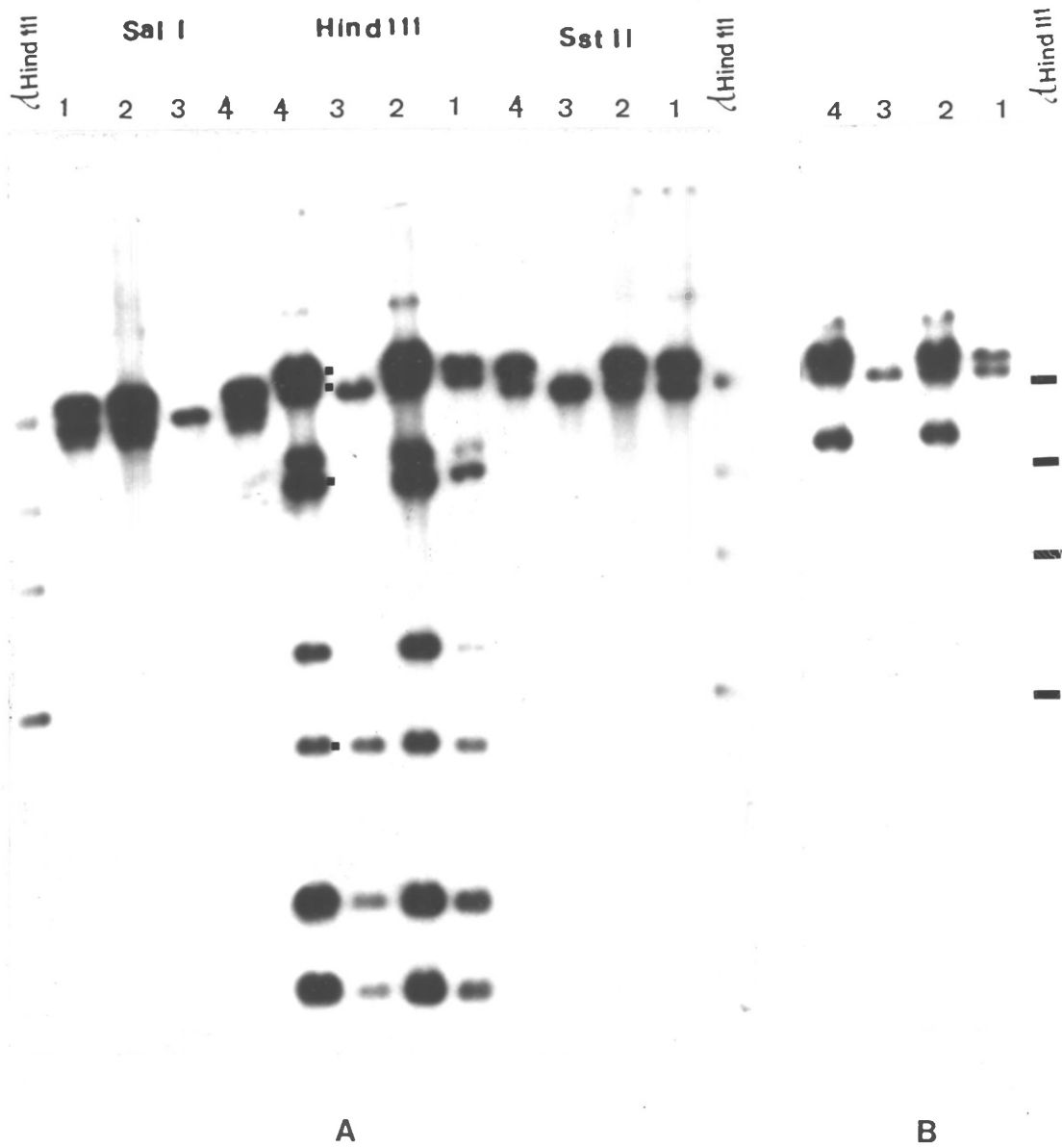


Figure 5.11 Southern blot of HindIII, SalI and SstII REs digested genomic DNA of S. marcescens::Tn7 probed with radiolabelled ColE1::Tn7 and pRS5 DNAs. Numbers 1, 2, 3 and 4 indicate four independent events of Tn7 insertion in S. marcescens.

A: DNA probed with radiolabelled ColE1::Tn7. The results shows for SalI and SstII cleaved DNA (lanes 1, 2 and 4), two Tn7-attTn7 flanking fragments whilst only one band of homology in one case (lane 3). HindIII RE cleaved DNA shows four Tn7-attTn7 flanking fragments in three cases [(1, 2 and 4), marked by dots]. These fragments are of approximate estimated sizes, 24 Kb, 22 Kb, 10 Kb and 3.5 Kb and are of the same sizes as seen in figure 5.9. Other HindIII fragment are the products of incomplete digestion. Only two HindIII RE Tn7-attTn7 flanking fragments in the case of number 3 (lane 3, marked by dots) are visible. This blot shows that Tn7 has inserted at two sites in three independent events (lanes 1, 2 and 4), and at one site in one independent event (lane 3).

B: HindIII RE digested genomic DNA probed with radiolabelled pRS5. The results show in three events (lanes 1, 2 and 4) two internal HindIII fragments of Tn7 and two of Tn7-attTn7 do not have any homology with the left end probe. An approximately 12 kb HindIII fragment in lanes 1, 2 and 4 which shows homology with pRS5 is due to partial digestion. The event number 3 (lane 3) has only one band of homology as expected. This blot shows that in all cases Tn7 insertion has occurred in a unique orientation.

The DNAs of these digests were probed with radiolabelled ColE1::Tn7 and pRS5 DNA and results are shown in figure 5.11. These results suggest the presence of two Tn7's in the chromosome of three independent events and one insertion of Tn7 in one event. In all these cases the Tn7 insertion was in one orientation as the HindIII RE digested genomic DNA have shown four Tn7 flanking fragments of homology with ColE1::Tn7 and only two with pRS5 probe.

It is only in the case of S. marcescens where two insertions of Tn7 were observed in 12 out of 13 events studied. By comparing the results of the Southern blots presented in figures 5.7, 5.8, 5.9 and 5.11 it is not possible to draw any strong conclusions about the insertional behaviour of Tn7 in S. marcescens. Therefore, instead of performing more Southern blots, a direct approach was adopted to clone the attTn7 sites from the genome of S. marcescens with two insertions of Tn7, for DNA sequence analysis (see chapter 6). However, the results gathered from the Southern blot analysis can be rationalised as follows.

In all independent cases of Tn7 insertions in the genome of S. marcescens, bacterial cells were treated with acridine orange following the conjugal transfer of RP4::Tn7 and pPH1JI. It may be possible that in step-1 or step-2 of this transfer the plasmids were not cured by acridine orange. Alternatively, S. marcescens may have some naturally occurring cryptic plasmid which had provided a

target site for Tn7. However, the bacterial cells should not have any plasmid after curing since the cells were sensitive to RP4 and pPH1JI encoded drug markers. Despite this, it does not answer the mystery concerning the orientation of Tn7 as shown in figure 5.8. One independent event (number 5 in figure 5.8 has shown three bands of homology with one end of the probe. The findings in this case suggest that the events of transposition may be still in progress in the bacterial cells. Only one independent insertion of Tn7 out of 13 was found at a specific site and in unique orientation. All remaining events have two Tn7's in S. marcescens with either no clarity about the orientation as shown in figure 5.8, or with one orientation as shown in figure 5.11.

It is possible to believe that Tn7 may have rearranged itself in the first experiment (see figure 5.8), thus yielding all four EcoRI RE fragments with plasmid pRS5 or right end Tn7. This DNA rearrangement may be induced by acridine orange. However, this rearrangement was not seen in the second experiment (see figure 5.11).

Tn7 has inserted in either one or two sites in S. marcescens but Tn7 specificity still needs to be clarified. Further evidence is presented in chapter 6 in favour of two attTn7 sites in the chromosome of S. marcescens. However, the Southern blot results indicate that in all except for one case, there are suggestions of

two Tn7's in S. marcescens. The possibility of a second transposition event during cell division can be considered to justify these results. This may indicate an unknown mechanism for selection of Tn7 transposition pathways in S. marcescens.

5.7 Tn7 has a Secondary attTn7 Site in the Chromosome of E. coli:

The final question addressed in this chapter was the ability of Tn7 transposition in an E. coli strain which had deletion in the primary attTn7 from the chromosome. The reason for this question was to determine the features of the recognitions involved in the Tn7 transposition selection process. Therefore any insertion into a secondary site in the absence of primary attTn7 would allow this site to subsequently be sub-cloned for a DNA sequence comparison to the attTn7 sites.

An E. coli strain C600 derivative $\Delta 7$ was constructed by Carlos Flores in this laboratory which has a deletion in the primary attTn7 site from the chromosome. This was achieved by first replacing a 4850 bp HindIII RE and BglIII RE fragment of plasmid pCPL6 (Lichtenstein and Brenner, 1982) with neomycin phosphotransferase gene (npt1) (Kn^r). The deletion of this fragment corresponds to the 350 bp to the right and 3500 bp to the left of the point of insertion of Tn7 in the primary attTn7 site. However this plasmid

still contains 500 bp further right and 1500 bp left to the deletion points.

This recombinant plasmid was then introduced to the bacterial strain C600 and after repeated selection for a plasmid incompatibility, it was possible to obtain a double reciprocal recombination event leading to a stable integration of npt1 in the chromosome. Therefore, the derived bacterial strain $\Delta 7$ has a kanamycin resistance gene in place of the attTn7 site and with glmS phenotype.

Bacterial strain $\Delta 7$ was utilised as a candidate for introducing Tn7 as a search for any secondary insertion of Tn7 in the chromosome. Tn7 was introduced in $\Delta 7$ via F' episome by mating with a bacterial strain 71/18 (F':Tn7). The choice of F':Tn7 was due to the fact that F' factor can mobilise the Tn7 and can be selected in the recipient cells easily. Another reason was that F' can be cured, by treating the cells with acridine orange, therefore the fate of Tn7 can be followed in the chromosome.

100 independent bacterial matings of $\Delta 7$ were performed with the 71/18 (F':Tn7) by tooth-picking the colonies onto non-selective YT plates and the recipient colonies were selected on LB plates containing Tp and Kn. After an overnight growth at 37 °C, the recipient colonies were tooth-picked onto another LB plate containing Kn and 100ug/ml acridine orange. This step was introduced to cure the F' and repeated at least 10 times by tooth-picking the

individual colonies onto the LB plates containing Kn and acridine orange.

The loss of F' factor was tested by the inability of bacterial cells to grow on a lawn of M13 derived phage, AJD18 which transduces the Ap^r. This vector was constructed for the selection of the loss of F' from the bacterial cells by cloning the plasmid pUC18 at the SmaI RE site of the M13 phage vector, mp10. After repeated curing of F', only seven colonies were unable to grow on this phage AJD18 lawn, thus unable to transduce Ap^r. The remaining 93 colonies were either able to grow onto the phage lawn, (suggesting that F' has not yet been lost) or have lost the Kn^r (suggesting that Tn7 has jumped into the Kn^r gene, thus inactivating the kanamycin resistance). By employing the strategy described above it was not easy to calculate the transposition frequency of Tn7.

However, the chromosomal DNAs of strain $\Delta 7$ colonies which had received the Tn7 and lost the F' factor were analysed to demonstrate the biochemical presence of Tn7 by the method of Southern (1975). The EcoRI RE digested genomic DNAs of the 7 bacterial $\Delta 7::Tn7$ colonies were probed with ³²[α P]-dCTP labelled ColEI::Tn7 and autoradiographed. The result of this blot is presented in figure 5.12 (A). These results show that Tn7 has inserted in all independent colonies of strain $\Delta 7$ at a unique site. In order to determine the orientation of Tn7 in $\Delta 7::Tn7$, DNA of four independent colonies was digested with HindIII RE

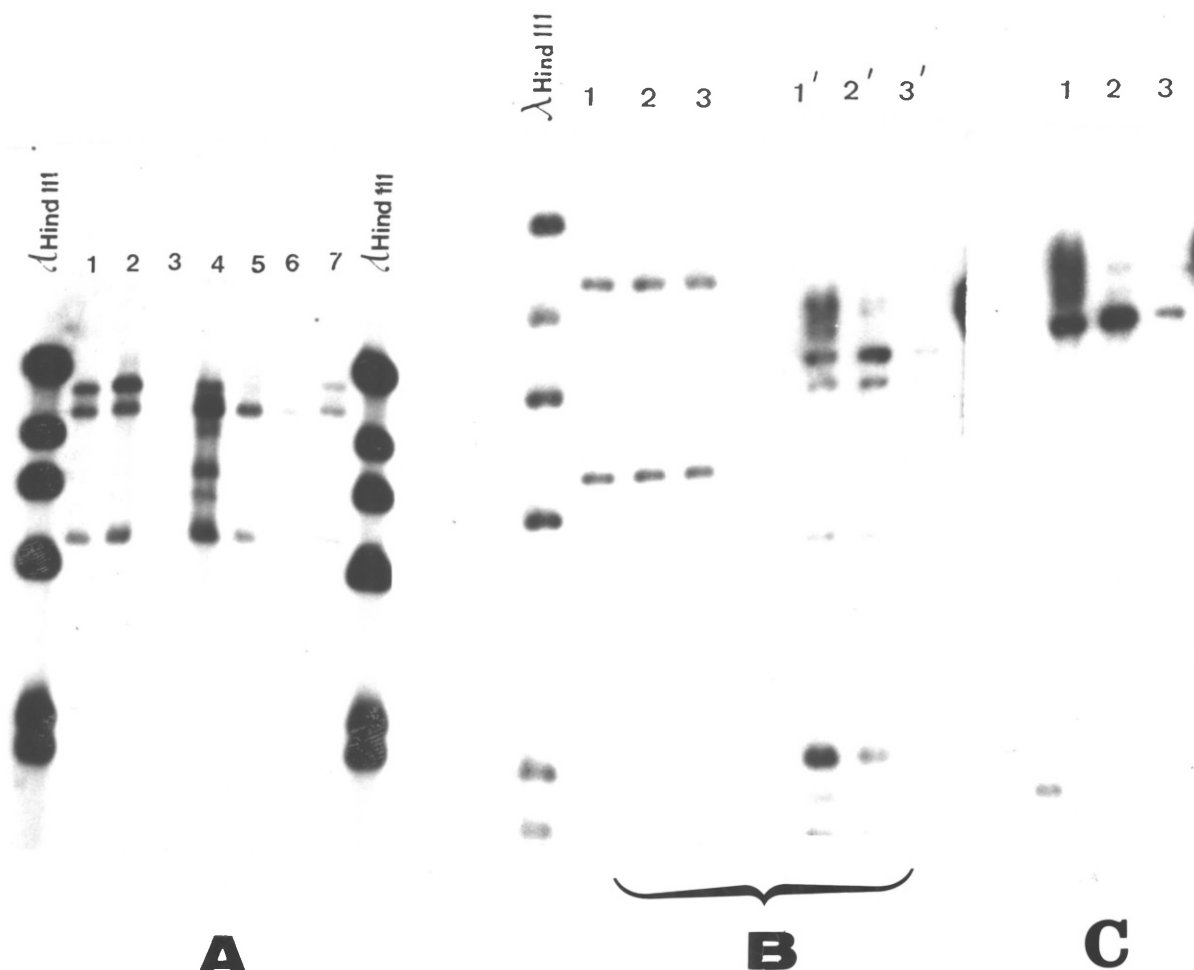


Figure 5.12 Southern blot of EcoRI, HindIII REs digested genomic DNA of $\Delta 7::Tn7$ probed separately with radiolabelled ColE1::Tn7, pFK::Tn7 and pRS5 DNAs.

A: lanes 1-7, seven independent insertions of Tn7 in the chromosome of $\Delta 7::Tn7$ cleaved with EcoRI RE and probed with radiolabelled ColE1::Tn7.

B: lanes 1-3 and 1'-3', three independent insertions of Tn7 (in duplicate) cleaved with EcoRI RE and HindIII RE respectively and probed with radiolabelled pFK::Tn7 DNA.

C: As B (1'-3') but probed with radiolabelled pRS5 DNA. It is clear in three cases the Tn7 insertion has occurred at a unique site in unique orientation; as two internal HindIII fragments of Tn7 and one Tn7-attTn7 did not show homology with the left end of Tn7. In lane 1, a 2 Kb fragment still showing homology with the left end probe of Tn7 suggests that it probably comes from the left end of Tn7 by the star activity of HindIII RE.

and probed with radiolabelled pFK7::Tn7 and pRS5 DNA (figure 5.12, B and C respectively). The result has suggested that in all four cases Tn7 has inserted in a unique orientation.

5.8 Experimental Conclusions:

The experiments described in this chapter have provided evidence for high frequency transposition of Tn7 in the chromosomes of A. tumefaciens, K. pneumoniae and S. marcescens. The transposition frequency of Tn7 in these bacterial species is approximately 90-96%. The frequency with which Tn7 is found to transpose between replicons is not influenced by the degree of selective pressure in its favour. Therefore the high frequency acquisition of Tn7 by the chromosome of several bacterial species (Towner et al., 1982) also seems to be the case in A. tumefaciens, K. pneumoniae and S. marcescens.

Southern blot results show that Tn7 insertion has occurred in a small region of A. tumefaciens and K. pneumoniae chromosomes in 9 independent derived strains. It seems that Tn7 has preferentially inserted at or near a single site which occurs once in the chromosome. This observation is in agreement with the previous reports of unique Tn7 insertion in the chromosome of E. coli (Lichtenstein and Brenner, 1981), P. aeruginosa (Caruso and Shapiro, 1982), C. crescentus (Ely, 1982) and X. campestris pv campestris

(Turner, 1984).

In contrast, Tn7 appears to have two specific attTn7 sites in the genome of S. marcescens. In all three bacterial species Tn7 insertion is found to occur in a unique orientation.

Another observation obtained from these experiments is that in E. coli, in the absence of primary attTn7 site, Tn7 inserts into a specific secondary attTn7 site. Southern blot results show that Tn7 has inserted into this site in a unique orientation like in other bacterial species. The frequency of transposition to the secondary attTn7 is not known.

CHAPTER 6

DNA Sequence Determination of the Tn7 Insertion Sites in the Chromosomes of A. tumefaciens, K. pneumoniae and S. marcescens

6.1 Introduction:

In the previous chapter, it was shown that Tn7 undergoes high frequency transposition into the chromosome of A. tumefaciens, K. pneumoniae and S. marcescens. It was also established, by Southern blot analysis, that Tn7 has unique insertion sites in the chromosome of A. tumefaciens and K. pneumoniae; however, Tn7 insertion in the chromosome of S. marcescens was found to be at two preferred sites.

The purpose of the experiments described in this chapter was to determine the DNA sequence of the target sites (attTn7) in the chromosome of these bacteria into which Tn7 transposes with very high frequency. This strategy was adopted to find any consensus DNA sequence at the insertion site of Tn7 which would provide a clue about the highly precise substrate/target site selection for high frequency transposition of Tn7.

To do this, it was necessary to clone Tn7 and Tn7 flanking DNA sequences from the chromosomes of these bacteria into suitable plasmid vectors, and then to sub-clone Tn7-attTn7

junction fragments into M13 phage vectors for DNA sequence analysis using Sanger's dideoxy chain termination method (Sanger et al., 1977).

6.2 Strategy for Cloning the Insertion Sites of Tn7 from the Chromosomes of Three Bacterial Species:

Tn7 encoded antibiotic resistance genes were used as dominant selectable markers for cloning large DNA restriction fragments containing Tn7 and attTn7 flanking chromosomal sequences from the three bacterial species into suitable plasmid vectors.

In a restriction map established by Gosti-Testu et al. (1983) Tn7 was shown to contain no restriction endonuclease sites for SalI, SstII, SmaI/XmaI or XhoI. This observation was exploited to clone the fragments containing the entire Tn7 element plus flanking sequences. Total genomic DNA from isolates of each species containing Tn7 in the chromosome was cleaved with SalI and ligated with suitable plasmid vectors. Selecting transformants for Tn7 encoded antibiotic resistance ensured the presence of the desired insert in recombinant plasmids.

In the cloning procedure, three plasmid vectors pUC18, pRK404 and pTJS75Kn were separately employed to obtain the recombinant clones. Figure 6.1 shows the restriction map and some related properties of each plasmid.

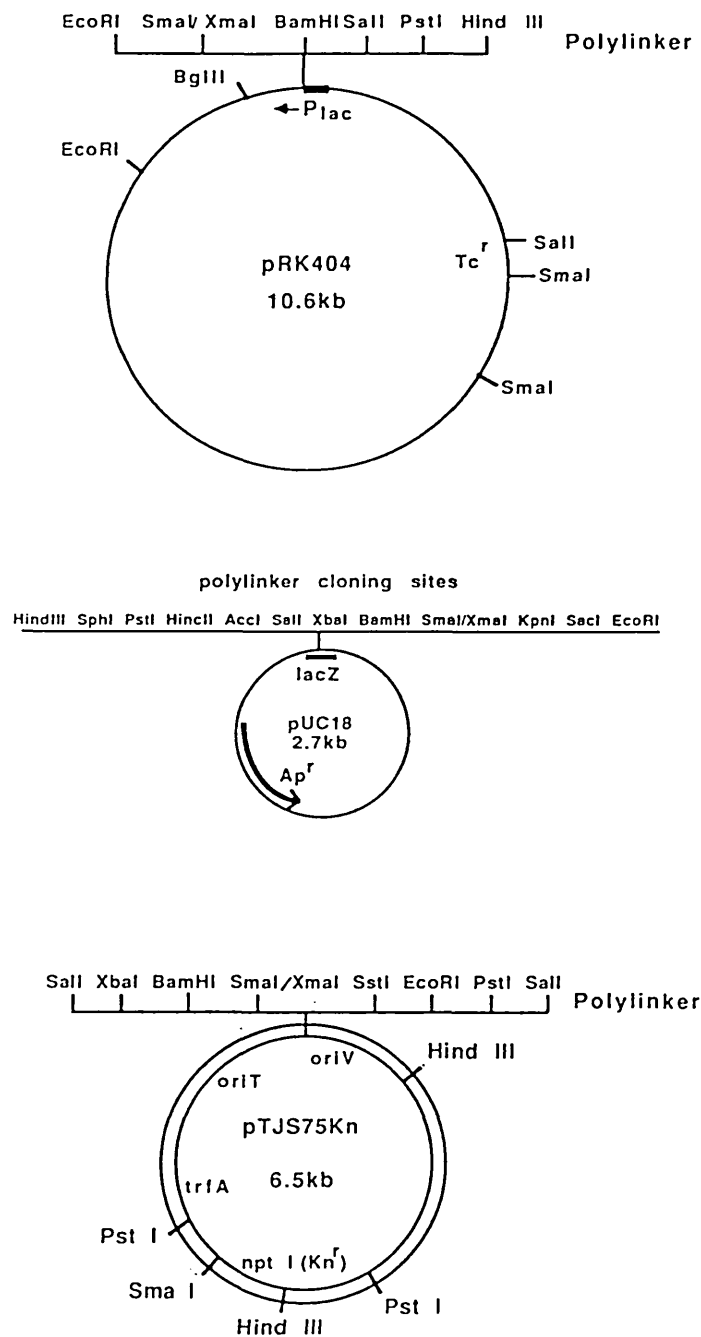


Figure 6.1 Restriction maps of plasmids pRK04, pUC18 and pTS75Kn.

The plasmid DNAs cleaved with SalI RE were dephosphorylated with calf intestinal phosphatase to prevent recircularisation of the vector during the ligation steps. After ligation of the SalI cleaved chromosomal DNA into the plasmid vectors described above, the DNA was transformed to E. coli cells DH5 and the colonies containing the putative recombinant clones were selected on MH plates containing trimethoprim and, when present, a plasmid encoded antibiotic resistance.

Possibly due to the large size of the insert in each case, pUC18 derived clones were found to be unstable and deletions of different sizes were observed in the recombinant clones (data not shown). The pRK404 derived recombinant plasmids were lost due to the absence of phenotypic selection for plasmid maintenance, as the only antibiotic resistance gene (Tc^r) was inactivated during cloning. Clearly, as Tn7 transposes with high frequency to the attTn7 site in the chromosome, Tn7-encoded drug markers cannot be used to select for stable inheritance of the pRK404::attTn7::Tn7 plasmid recombinant.

The reasons for choosing the third plasmid pTJS75Kn for cloning Tn7 and attTn7 flanking sequence were three fold; (i) it provides a polylinker for cloning purposes, (ii) it can be selected for by (Kn^r) after cloning the insert, and (iii) it can accommodate large inserts as it is an RP4 derivative; RP4 is a large low copy number plasmid.

The ligation mixture was added to competent DH5 cells and the pTJS75Kn derived recombinant transformants were selected on MH plates containing trimethoprim and kanamycin. Stable recombinant plasmids were obtained in all cases except for A. tumefaciens. In the following section, a brief summary of the plasmid clones obtained containing Tn7 and attTn7 sites from each bacterial chromosome is described.

6.2.1 Cloning of attTn7 Site from the Chromosome of A. tumefaciens:

Several attempts were made to clone the SalI RE fragment (20 Kb, $\pm$ 0.5Kb) containing Tn7 and the flanking DNA sequence from two isolates of A. tumefaciens into plasmid pTJS75Kn, but all failed to yield the desired recombinant plasmids.

Two clones were found after selection on kanamycin and trimethoprim plates which had apparently suffered large deletions of approximately 14-15 Kb. The clones were named pAIQ1:: Δ Tn7 and pAIQ2:: Δ Tn7 and when cleaved with SalI, HindIII and EcoRI REs showed the deletion of a large region (data not shown).

The host cells used to propagate this recombinant plasmid were recA⁻, therefore the deletion event in both clones must be independent of the product of the recA gene. Subsequent attempts made recently in this laboratory to

clone the virgin attTn7 site from the chromosome of A. tumefaciens in a cosmid vector also failed to yield any clones (Bun Ng, personal communication).

The reasons for the inability to clone this piece of A. tumefaciens DNA in E. coli remain obscure. One possibility is that this DNA encodes a toxic gene product; toxic despite the low copy gene copy number allowed by the plasmid vector. Another possibility is that the DNA may have some secondary structure which gives rise to deletions in E. coli.

As the precise localisation of these deletions was not clear, it was possible that the attTn7-Tn7 junctions remained and the plasmid recombinants were used to sub-clone Sau3AI RE fragments into M13, mp10 for the purposes of DNA sequence analysis (see section 6.3.1).

6.2.2 Cloning of the attTn7 Sites from the Chromosome of K. pneumoniae:

The SalI RE fragment containing Tn7 and the flanking chromosomal DNA from 6 independent isolates of K. pneumoniae were successfully cloned into plasmid pTJS75Kn. The recombinant plasmids of each isolate were named pKIQ1::Tn7 to pKIQ6::Tn7. From the previous chapter (figure 5.10) it was seen that Tn7 is encompassed in a 21 Kb $\pm$ 0.5 Kb SalI RE fragment of K. pneumoniae chromosomal

DNA. When all six plasmids (pKIQ1::Tn7-pKIQ6::Tn7) were cleaved with SalI RE, they released this restriction fragment. These plasmids were also cleaved with a battery of other REs. In all cases, stable maintenance of plasmid and insert containing Tn7 was observed. All recombinant plasmids were used for the construction of Sau3AI libraries in the M13 phage vector, mp10, to determine the DNA sequence of the attTn7 site. The sub-cloning sequencing strategy is discussed in section 6.3.2.

6.2.3 Cloning of the attTn7 Sites from the Chromosome of S. marcescens:

It was established by Southern blots in chapter 5 that Tn7 had transposed in 13 out of 14 cases into two sites in S. marcescens as shown in figures 5.7, 5.9, 5.10 and 5.11. In order to clone the two attTn7 sites from S. marcescens, chromosomal DNA with two insertions of Tn7 (see figure 5.11, lane 2), was cleaved with SalI RE and ligated into the SalI RE site of pTJS75Kn. The ligated DNA was used to transform DH5 cells as described previously. Twenty transformant colonies showing resistance to trimethoprim and kanamycin were obtained. Recombinant plasmid DNA was isolated from each transformant as described in chapter 2, section 2.2.2B(ii). Plasmid DNA was digested with SalI RE and fractionated on a 0.8% agarose gel (see figure 6.2). Two size classes of insert were either 20 Kb or 25 Kb ($\pm$ 0.5 Kb). Since the Southern blot results described in

chapter 5 (figure 5.10) showed that the two SalI RE fragments containing Tn7 in the genomic DNA were of the approximate sizes of 20 and 25 Kb ($\pm$ 0.5 Kb), this suggested that both Tn7 insertion events had been cloned and no internal deletions had occurred in any of the recombinant clones.

In chapter 5 it was discussed that the second Tn7 insertion site may not be in the S. marcescens chromosome but rather be retained on the plasmid donor RP4::Tn7 or have transposed to the pPH1JI. Drug markers for these plasmids had been lost but the plasmids could still have been present in a rearranged form.

As one of the SalI RE fragments of RP4 is 12 Kb and if Tn7 had inserted into this fragment in the RP4::Tn7 donor it would produce a 25 Kb fragment of Tn7 plus attTn7. It was possible that one of these fragments was accidentally cloned in plasmid pTJS75Kn. The same sized SalI RE fragment was observed in the genomic Southern blots (as shown in chapter 5, figures 5.10 and 5.11) and in the recombinant plasmids. Since RP4 and pPH1JI have extensive DNA homology, RP4 was used as a hybridisation probe against SalI RE digests of the 12 recombinant plasmids, to rule out the possibility that an insert in recombinant clones came from either plasmid. The RP4 plasmid DNA was a gift from Stan Theophilus, Imperial College, London.

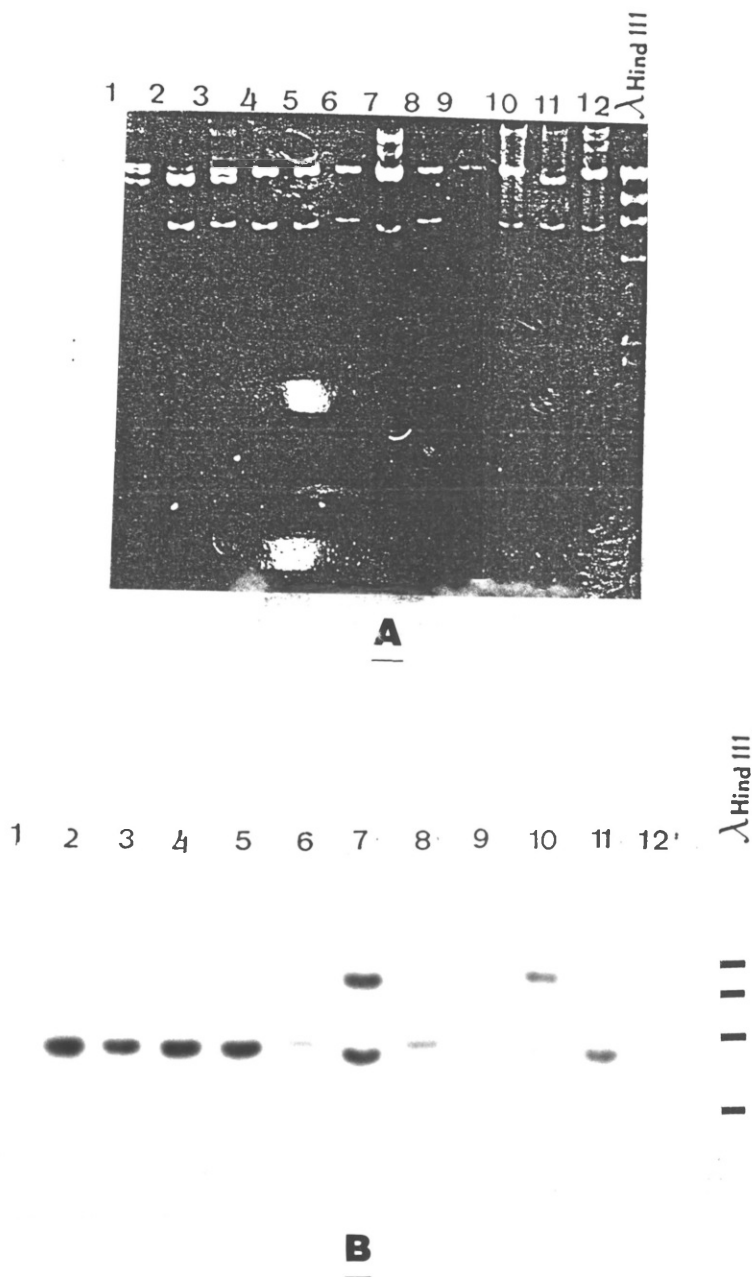


Figure 6.2 Photograph and Southern blot of SalI RE digested pTJS75Kan derived recombinant plasmids (pSIQ1::Tn7-pSIQ12::Tn7) probed with radiolabelled RP4 DNA.

A: Photograph of 0.8% agarose gel.

B: Southern blot of A. In most cases the upper fragment containing Tn7-attTn7 does not show homology with RP4 DNA. In lanes number 3, 7 and 10 the additional signal of homology with the larger SalI fragment is due to incomplete digestion of recombinant plasmids. The positive control of this experiment is plasmid pTJS75Kn.

From the results shown in figure 6.2 it can be seen that plasmid RP4 does not display any homology with the insert of (pSIQ1::Tn7-pSIQ12::Tn7). Since the vector pTJS75Kn has homology with RP4, in some cases there is a weak signal of homology caused by the incomplete digestion of recombinant plasmids. This suggests that recombinant plasmids pSIQ1::Tn7-pSIQ12::Tn7 contain inserts which did not originate from RP4 or pPH1JI.

Therefore it can be concluded that the recombinant clones either have inserts of Tn7 and Tn7 flanking sequence from the chromosome or from a cryptic plasmid of S. marcescens. Representatives of the plasmid recombinants containing a 20 Kb or 25 Kb insert were selected and cleaved with different REs, plasmid number 4 produced a different restriction pattern to numbers 2 and 3. These results confirmed that both SalI RE fragments are represented in the recombinant clones. Plasmid number 4 has a SalI insert of approximately 25 Kb ($\pm$ 0.5 Kb) and this insert corresponds to the upper SalI RE fragment in genomic blots as shown in figures 5.10 and 5.11. This plasmid was designated as pSIQI::Tn7. The other plasmids which have the SalI insert of approximately 20 Kb ($\pm$ 0.5 Kb) were designated as pSIQII::Tn7. Both pSIQI::Tn7 and pSIQII::Tn7 are stably maintained in E. coli strain DH5.

In order to confirm that both plasmids have the insert from S. marcescens, nick translated DNA from pSIQI::Tn7 and

pSIQII::Tn7 was used as a hybridisation probe against Tn7 free S. marcescens colonies. Both probes show homology with S. marcescens total DNA, indicating that each plasmid contains DNA from S. marcescens (data not shown). The probes made from either Tn7 or the plasmid pTJS75Kn did not show any homology with the colonies of S. marcescens.

DNA from each plasmid pSIQI::Tn7 and pSIQII::Tn7 was used to construct the Sau3AI libraries in M13 phage vector, mp10 for determining the DNA sequence of Tn7 and attTn7 junction fragments as discussed in section 6.3.3 below.

6.3 DNA Sequence Analysis of attTn7 from the Chromosome of the Three Bacterial Strains:

It was decided to construct Sau3AI libraries of each plasmid discussed above, in the M13 phage vector, mp10. By doing this it was possible to identify the left and right junction fragments from the recombinant plasmids containing Tn7 and attTn7 DNA, using Tn7-attTn7 junction fragments as probes.

6.3.1. The Construction of Sau3AI Library of pAIQ1:: Δ Tn7; pKIQ1::Tn7-pKIQ6::Tn7 and pSIQI::Tn7/pSIQII::Tn7 in M13, mp10:

DNA from each plasmid was digested with Sau3AI RE and ligated with BamHI-digested mp10 (RF) DNA. Sau3AI and

BamHI share the same tetranucleotide (5'-GATC-3') in their recognition sites and produce compatible "sticky ends". The ligated site is a hybrid Sau3AI/BamHI RE site and is unlikely to serve as a substrate for BamHI enzyme. Thus, it was possible to redigest the ligated molecules with BamHI to select against the recircularised vector molecules without any insert; linear molecules transform E. coli less efficiently than circular molecules.

The DNA was used to transform E. coli strain TG2, which was then plated with a lawn of M13 sensitive TG2 onto 2X YT plates with IPTG and X-gal [as described in chapter 2, section 2.2.2E(i)]. In order to identify the left and right Sau3AI RE Tn7/attTn7 fragments, 1000-2000 colourless M13 plaques were transferred by tooth-picking in a gridded array onto a fresh lawn of TG2 containing IPTG and X-gal. Plaques of M13, mp10 (negative control) and the M13 derived clones Pst9 and OE6 (positive controls) were included. After overnight incubation phage DNA from the resultant plaques were transferred to nitrocellulose filters in duplicate.

6.3.2 Identification of Right End Junction Fragments from the Sau3AI Library of pAIQ1:: Δ Tn7; pKIQ1::Tn7-pKIQ6::Tn7 and pSIQI::Tn7/pSIQII::Tn7:

For the identification of the Sau3AI right-end junction fragment, filters from each M13 library were hybridised

with nick translated 925 bp PstI RE fragment from the M13 derived clone, Pst9. This fragment was isolated from a 0.8% agarose gel [see chapter 2, section 2.2.2B(vi)] and nick translated using radioactive [³²P]-dCTP as described in chapter 2, section 2.2.2D(ii). The PstI RE fragment contains 389 bp of the right end of Tn7 and the 536 bp of attTn7 from the E. coli chromosome.

M13 Sau3AI libraries from each plasmid were screened as described in chapter 2, section 2.2.2F(ii). Plaques showing homology with this fragment should have the right end of Tn7 and attTn7 site. The PstI RE fragment which was used as probe contains two additional Sau3AI RE sites from the right end of Tn7, so the clones identified with this probe may not necessarily contain the Sau3AI fragment that includes the attTn7 sequence.

Therefore, to find the clones containing the right junction fragment and to define the orientation, single-stranded DNA of all positive clones was annealed in vitro with the single-stranded DNA of M13 derived phages, 31 and 34. These vectors contain a small insert of 41 bp of the right end of Tn7 with 389 bp of attTn7 from the E. coli chromosome, one in each orientation, with reference to the universal priming site of the M13 phage vector (see figure 3.3 for description).

Positive clones having regions complementary to either 31 or 34 showed retarded mobility upon electrophoresis in

agarose gel [see chapter 2, section 2.2.2E(ii)]. These clones were named on the basis of the orientation of the insert. Clones which hybridised to 31 were said to be in orientation 1 and those that hybridised to 34 in orientation 2. By employing this strategy the right end junction clones from pKIQ1::Tn7-pKIQ6::Tn7 and pSIQI::Tn7/pSIQII::Tn7 were identified. None were obtained from pAIQ1::ΔTn7 or pAIQ2::ΔTn7. All identified clones were saved for DNA sequence analysis.

6.3.3 Identification of the Left End Junction Fragment from the Sau3AI Library of pAIQ1:: ΔTn7; PKIQ1::Tn7-pKIQ6::Tn7 and pSIQI::Tn7/pSIQII::Tn7:

M13 based Sau3AI libraries from each plasmid were screened as before to identify left-end junction fragments using a nick translated probe of the 307 bp EcoRI RE fragment from the M13 phage vector, OE6. This clone is already described in chapter 3 (figure 3.3) and the EcoRI insert contains the 210 bp of the left end of Tn7 and 97 bp of the left end of the attTn7 from the E. coli chromosome. The single-stranded DNA of all positive plaques showing homology with this fragment were then annealed with the single-stranded DNA of OE6 as described in chapter 2, section 2.2.2E(ii). Of the clones that hybridise to the nick translated probe, those that also hybridise to the single-stranded OE6 DNA were said to be in orientation 1 and those that do not anneal were said to be in orientation 2.

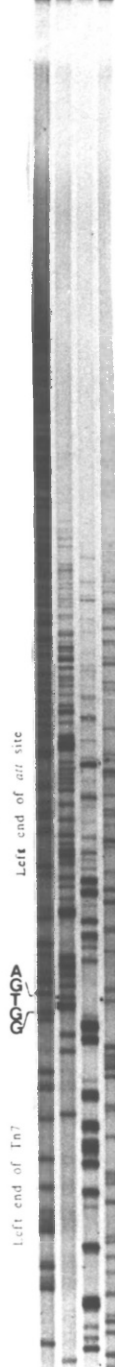
6.4 DNA Sequencing Analysis of the Right and Left Junction Fragments of Tn7-attTn7:

Single-stranded DNA from the right junction fragments, (RJK1/RJK2, RJS1.I and RJS2.II) and the left junction fragments (LJA2, LJK1/LJK2, LJS2.I and LJS2.II) was isolated and used in DNA sequencing reactions as described in chapter 2, section 2.2.2D(viii). Figure 6.3 shows the autoradiograph of DNA sequencing gels of left and right end of Tn7-attTn7 fragments in these M13 clones. The DNA sequence from the left and right junction fragments were obtained from many gel readings and are compiled together in figures 6.4, 6.5 and 6.6. The DNA sequence at the attTn7 sites is taken after removing the DNA sequence of left and right ends of Tn7 in the junction fragments. However, the target duplication sequence is considered only once in the final sequence. Since the right end junction fragment from A. tumefaciens could not be obtained due to the large deletion in plasmid pAIQ:: Δ Tn7, the sequence presented in figure 6.4 is only from the left end junction fragment. The DNA sequence presented in figures 6.4, 6.5 and 6.6 is the most accurate consensus sequence, since some gels have some ambiguities. The 5 bp sequence at the target which was duplicated upon insertion of Tn7 in the chromosome is shown only once. In A. tumefaciens (pAIQ1:: Δ Tn7/pAIQ2:: Δ Tn7) the 5 bp target sequence was (5'-GGTGA-3') which was obtained by reading the left junction fragment (LJA2).

Figure 6.3 DNA sequencing gels of Tn7-attTn7 (left and right) junction fragments from A. tumefaciens::Tn7, K. pneumoniae::Tn7 and S. marcescens::Tn7 (both sites). The right end junction fragment from A. tumefaciens::Tn7 was never obtained. S. marcescens has two attTn7 sites as marked by I and II. The point of insertion of Tn7 in each junction fragment from each species is shown by dots.

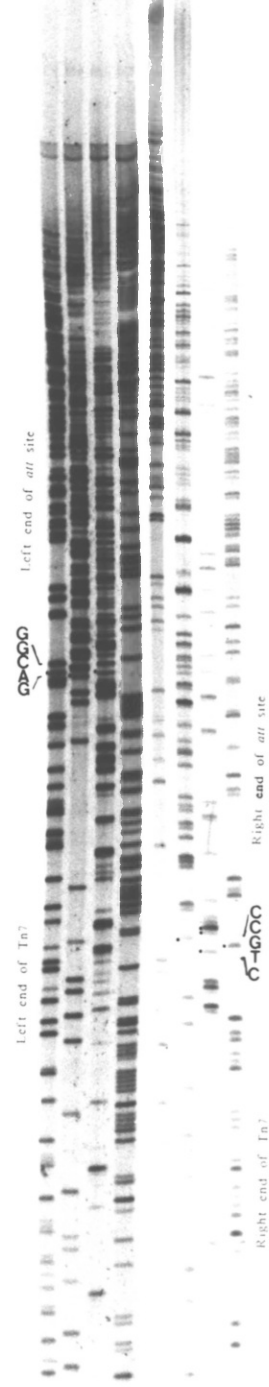
Agrobacterium tumefaciens

LJA1
AGCT



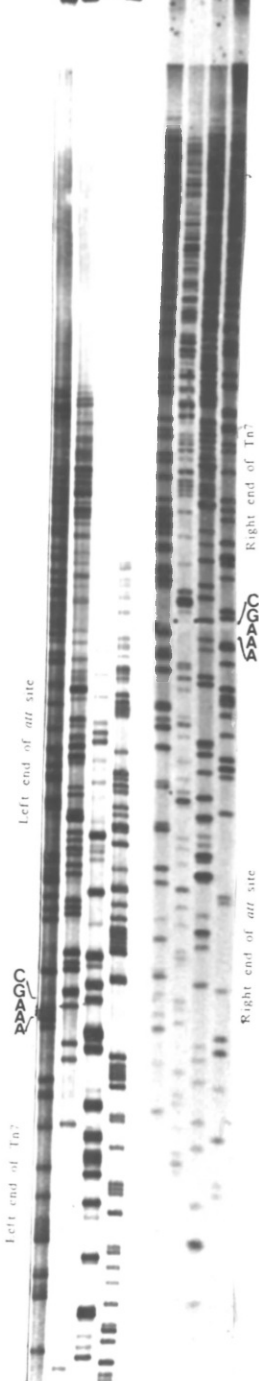
Klebsiella pneumoniae

LJK2 RJK2
AGCT AGCT



Serratia marcescens

LJS2.1 RJS1.1
AGCT AGCT



Serratia marcescens

LJS2.11 RJS2.11
AGCT AGCT

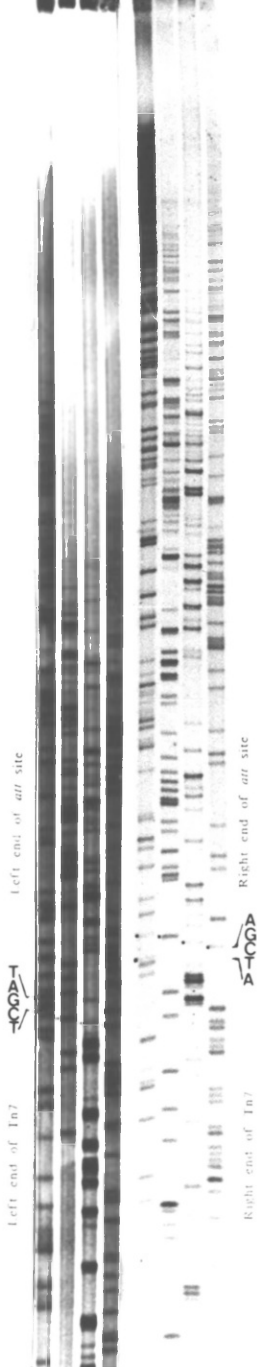


Figure 6.4 DNA sequence of attTn7 sites of A. tumefaciens and K. pneumoniae. These sequences were obtained from the plasmids pAIQ1::ΔTn7 and pKIQ1::Tn7-pKIQ6::Tn7 respectively. The point of insertion of Tn7 in each case is displayed in the box. The star is the translational stop codon of DNA encoding the identical amino acids at the carboxy-terminal of the glmS of E. coli. The DNA with potential to form cruciform structure is shown by its ΔG value.

Agrobacterium tumefaciens (incomplete attTn7 sequence)

```

      10
5'-GATCGCCTCT TCATCACC-3'
3'-CTAGCGGAGA AGTAGTGG-5'

```

Klebsiella pneumoniae (attTn7 sequence)

```

      10      20      30      40      50
5'-GATCGGCGCC TGGCTCACCG TGGTGGACTG GTTCGCCAGC CGGAGGAGCG
3'-CTAGCCGCGG ACCGAGTGGC ACCACCTGAC CAAGCGGTCTG GCCTCCTCGC

```

```

      60      70      80      90      100
ACGAGGCGCT GCCGTTGCGC GATTGACGGC GAGCGCCGGG GCCGCATGCC
TGCTCCGCGA CGGCAACGGC CTAAGTGGCG CTCGCGGCCC CGGCGTACGG

```

ΔG = -27.0

```

      110      120      130      140      150
TGCCTGCCCC GGCTTCTCGG GCACGCGGCC GGCGTTACTC GACGGTGACC
ACGGACGGGG CCGAAGAGCC CGTGCGCCCG CCGCAATGAG CTGCCACTGG

```

ΔG = -49.6

```

      160      170      180
GACTTGCCCA GGTTGCGCGG CTGGTTCGACG GATC-3'
CCTAACCGGT CCAACGCGCC GACCAGCTGC CTAG-5'
- glucosamine synthetase -----

```

Figure 6.5 DNA sequence of attTn7 (1st site) of S. marcescens. This sequence was obtained from the plasmid pSIQI::Tn7. The 5 bp point of insertion of Tn7 is displayed in the box. The star is the translational stop codon of DNA encoding the identical amino acids at the carboxy-terminal of glmS of E. coli. The DNA with potential to form cruciform structure is shown by its ΔG value.

```

      10           20           30           40           50
5'-GATCTGTGAA CAGCACATTC CCAAGCAATA AAAACCGCTT GCGCGCCCTA
3'-CTAGACACTT GTCGTGTAAG GGTTCGTTAT TTTTGGCGAA CGCGCGGGAT

```

```

      60           70           80           90          100
TCCGCTTGAA TATGACAACG TTTATCACCG AAACGGGCAGC GCTGTCATCT
AGGCGAACTT ATACTGTTGC AAATAGTGGC TTTGCCGTCG CGACAGTAGA

```

$\Delta G = -8.4$

```

      110          120          130          140          150
TTAAGGTCAC TCATGTTAAA ACGGCGCTAA GGCCGGCTTT TGAAGTTTAC
AATTCCAGTG AGTACAATTT TGCCGCGATT CCGGCCGAAA ACTTCAAATG

```

$\Delta G = -8.4$

```

      160          170          180          190          200
CTGCGACAAA TTA CTCTACG GTGACAGACT TCGCCAGGTT GCGCGGCTGG
GACGCTGTTT AATGAGATGC CACTGTCTGA AGCGGTCCAA CGCGCCGACC

```

* ----- glucosamine synthetase -----

```

      210          220
TCAACGTCGG TGCCCTTGAT C-3'
AGTTGCAGCC ACGGGA ACTA G-5'
-----

```

Figure 6.6 DNA sequence of attTn7 (2nd site) of S.marcescens. This sequence was obtained from the plasmid pSIQII::Tn7. The 5 bp point of insertion of Tn7 is displayed in the box. The dotted region is homologous with the MetG tyrosyl tRNA synthetase gene of E. coli. The DNA with potential to form cruciform structure is shown by its ΔG value.

```

      10      20      30      40      50
5'-GATCTGTTCC GTTCACATCA GACAAAAAAT GGAAGTGCAC CACAGCAAGC
3'-CTAGACAAGG CAAGTGTAGT CTGTTTTTTT CCTTGACGTG GTGTCGTTCG

      60      70      80      90      100
GGCGCTCGCA TGGCTTTCCC GCGCCGCCGA GGTGGGTCTA TTGGAAAGTG
CCGCGAGCGT ACCGAAAGGG CGCGGCGGCT CCACCCAGAT AACCTTTCAC

      110      120      130      140      150
AGATACAATC AATCTGTTAT CGACTCACCA CCGGTGATGA CAGTGAACGG
TCTATGTTAG TTAGACAATA GCTGAGTGGT GGCCACTACT GTCACTTGCC
                        . . . . .
                        ΔG = -3.2

      160      170      180      190      200
CGCGCCGATG TCAAAGTACA AAAGAGAATA GGATGGACGC TTTTGGCATC
GCGCGGCTAC AGTTTCATGT TTTCTCTTAT CCTACCTGCG AAAACCGTAG
..      . . . . .

      210      220      230
AAGCCTTTAT CTCCTTTACC TGGCTGTTAT AACGATC-3'
TTCGGAATA GAGGAAATGG ACCGACAATA TTGCTAG-5'

```

<u>E. coli</u>	U E V T V K A L N R P Q D R R H R Q D
<u>K. pneumoniae</u>	U E V T V K A L N R P Q D V S
<u>S. marcescens</u>	U E V T V K A L N R P Q D V D T G K I

Figure 6.7 Predicted amino acid sequence at the C-terminal end of glmS coding region in the attTn7 sites of E. coli, K. pneumoniae and S. marcescens (1st site). The E.coli attTn7 sequence is taken from Lichtenstein and Brenner (1982) and Walker et al. (1984). The U in each case is the end at the C-terminal region of the polypeptide.

Table 6.1 One letter amino acid code.

One- and Three-Letter Amino Acid Abbreviations

A = Ala	= Alanine
C = Cys	= Cysteine
D = Asp	= Aspartic acid
E = Glu	= Glutamic acid
F = Phe	= Phenylalanine
G = Gly	= Glycine
H = His	= Histidine
I = Ile	= Isoleucine
K = Lys	= Lysine
L = Leu	= Leucine
M = Met	= Methionine
N = Asn	= Asparagine
P = Pro	= Proline
Q = Gln	= Glutamine
R = Arg	= Arginine
S = Ser	= Serine
T = Thr	= Threonine
V = Val	= Valine
W = Trp	= Tryptophan
Y = Tyr	= Tyrosine
B = Asx	= Asp or Asn, not distinguished
Z = Glx	= Glu or Gln, not distinguished, or pyrrolidone carboxylic acid
X = X	= Undetermined or atypical amino acid

- 1) 36 bp box in glmS gene
- 2) 2 or 3 helical turn spacing between end of 36bp box & beginning of 5bp target
- 3) Role of hairpins
- 4) Role of the GCNGG motif

Figure 6.8 Comparison of chromosomal attTn7 sites of E. coli (Ec), K. pneumoniae (Kp) and S. marcescens (Sm) 1st site. Positions where identical nucleotides occur in the coding region of the glmS are shown in shaded boxes. The asterisk shows the translational stop codon of the glmS gene. In all cases DNA encoding the C-terminal domain of glmS in the 70 bp attTn7 is 100% conserved. $\Delta R25$ and $\Delta L849$ are the limits of right and left end deletions respectively (see chapter 4, section 4.3). The Tn7 5 bp target duplication sequence (5'-GCNGG-3') is marked in a box. In E. coli, K. pneumoniae and S. marcescens the occurrence of a 5 bp sequence motif (5'-GCNGG-3') is underlined.

6.5 Analysis of attTn7 Sites:

A detailed comparison of attTn7 sites from the chromosomes of K. pneumoniae and S. marcescens with the previously sequenced E. coli attTn7 site (Lichtenstein and Brenner, 1982), led to the following observations.

Analysis of DNA sequence of six independent insertions of Tn7 in the chromosome of K. pneumoniae indicate that in all cases Tn7 has inserted at an identical position generating the same 5 bp target duplication 5'-GCAGG-3'. Thus, as with the E. coli attTn7 site, Tn7 transposition to this K. pneumoniae site is very precise.

In K. pneumoniae and S. marcescens (1st site), Tn7 has inserted at the 3' end of a sequence which is very similar to the right end of the E. coli attTn7 (i.e. the carboxy-terminal of glmS gene). Thus a comparison of these attTn7 sites in the DNA sequence up to the termination codon of the glmS (which includes the 36 bp glmS box) shows homology of; 29/36 bp for K. pneumoniae versus E. coli, 28/36 bp for S. marcescens (1st attTn7) versus E. coli, and 33/36 bp for S. marcescens (1st attTn7) versus K. pneumoniae. The predicted amino acid sequence from this region is 100% identical, therefore suggesting in all cases Tn7 insertion has occurred at the 3' end of the carboxy-end of the glmS gene.

This high degree of DNA sequence homology in this region between these three bacterial strains, together with the

data obtained in the deletion analysis of the E. coli attTn7 site, argues strongly that this 36 bp sequence is an important functional domain required by Tn7 for high frequency/specificity target site selection.

While the DNA sequence consensus is striking and could account for the unique occurrence of attTn7 sites in the chromosomes of these bacteria, it is not the actual site of Tn7 insertion.

A comparison of the DNA sequences surrounding the actual insertion sites indicates that there is no clearly conserved consensus sequence. However, the sequence 5'-GCN GG-3' is present in the 5 base pair target duplication in E. coli and K. pneumoniae and occurs several other times in the K. pneumoniae attTn7 site (see figure 6.8). Although the 5'-GCN GG-3' motif is not within the 5 bp target duplication of the S. marcescens attTn7 site (1st site) it is present close-by (see table 6.2). The incomplete sequence of attTn7 from A. tumefaciens does not show such a sequence motif. Similarly, in the second attTn7 site from S. marcescens, the DNA sequence does not have any convincing homology with the attTn7 sites of E. coli, K. pneumoniae and S. marcescens (1st site). The 5'-GCN GG-3' sequence motif in three similar attTn7 sites may constitute a second and separate functional domain required for target site selection. The DNA sequence between these two putative functional domains is not conserved in the E. coli and K. pneumoniae attTn7 sites and in the S.

marcescens attTn7 site (1st site).

The distance between the last nucleotide of the 36 bp conserved domain and the first nucleotide of the 5 bp target duplication is 20 bp, 31 bp and 20 bp in E. coli, K. pneumoniae and S. marcescens (1st site) respectively. This spacing allows two (E. coli and S. marcescens) or three (K. pneumoniae) helical turns of the DNA. Thus in the S. marcescens (1st site) the 5'-GCNCG-3' motif is present 23 bp from the 36 bp glmS box, a non-integral (2.19) number of turns away. Interestingly, in S. marcescens (unlike E. coli and K. pneumoniae) this sequence motif is not the 5 bp target duplication (see table 6.2).

The spacing between the glmS box and the point of insertion of Tn7 in these attTn7 sites, being an integral number of helical turns away, suggests that DNA sequences recognised by Tn7 and/or host-encoded proteins during target site selection would face on the same side of the helix. Ptashne (1986) has proposed a looping out model to explain how non-adjacent DNA sequences are brought together via protein-protein interactions to regulate gene expression. In this model one protein binds to one site, another protein to the other site. Then protein-protein interactions result in the looping out of the intervening DNA. When the two DNA binding domains are close together, in order for the protein-protein interactions required for function to take place, the proteins must bind to the same side of the DNA double helix.

Table 6.2 A comparison of Tn7 target duplication in attTn7 sites and non-attTn7 sites.

Target	5 bp target duplication and 5 bp GCNCG motif	Number of independent studies	Reference
<u>attTn7</u> site			
<u>E. coli</u>	5'- GCGGG -3'	2	Lichtenstein and Brenner, (1982)
<u>K. pneumoniae</u>	5'- GCAGG -3'	6	This study
<u>S. marcescens</u> (1st <u>attTn7</u>)	5'- AAAGCGGG -3'	1	This study
<u>S. marcescens</u> (2nd <u>attTn7</u>)	5'- TCGAT -3'	1	This study
<u>A. tumefaciens</u>	5'- GGTGA -3'	1	This study
non- <u>attTn7</u> site			
ColE1	5'- TCTTC -3'	1	Lichtenstein and Brenner, (1982)
R68	5'- GTATT -3'	2	Nash and Krishnapillai, (1987)
R68	5'- GACAC -3'	1	"

The 5 bp sequence (5'-GCNCG-3') motif is shown in bold type in attTn7 sites and the 5 bp target duplication is shown in a box.

This requires a spacing between the two domains of an integral number of helical turns. For example, for efficient repression of the gal operon of E. coli (Irani et al., 1983), transcription of SV40 early genes (Takahashi et al., 1985) and maintenance of lambda lysogeny (Hochschild and Ptashne, 1986), the interaction of two proteins bound to two sites plays a vital role in regulation of gene expression. This can only be achieved if proteins are located on the same side of the DNA helix. It has been shown that introduction of half integral turns, so placing

the proteins on the opposite sides of the helix, reduce the gene expression. Energy would be required to bend or twist the DNA for proper alignment.

In the case of attTn7 sites it may be two proteins bound at two positions (separated by 2 or 3 helical turns) which interact by simple looping out of the intervening DNA for efficient (high frequency) site-specific transposition of Tn7. In order to test this, further experiments were performed and are described in chapter 7.

Another common feature in the structural organisation of attTn7 sites was the potential of DNA to form strong secondary structures downstream of the region of homology in the DNA encoding the C-terminal region of glmS. The transposon Tn7 has inserted close to these potential secondary structures in all cases. The thermodynamic values (Kcal per mol) of each potential cruciform are listed below the DNA sequence of the attTn7 sites (see figures 6.4, 6.5 and 6.6). These values were calculated according to the rule of Tinoco et al. (1973). Whether or not these secondary structures within the attTn7 sites are required for high frequency transposition of Tn7 are addressed in chapter 7.

The insertion of site-specific transposons Tn1545 and ISH1 have also been shown to occur in a DNA sequence which has potential to form cruciform structure (see table 1.1) (Caillaud and Courvalin, 1987; Simsek et al., 1982).

It has been shown for crucial biological processes such as transcription, replication and recombination that efficient enzymatic manipulation of DNA requires the presence of negative superhelical turns in the substrate DNA (see review by Lilley, 1984). The superhelical stress encourages the DNA to assume a wide variety of unusual structures such as cruciforms, slippage loops and left-handed Z DNA. The presence of potential cruciforms within attTn7 suggests that they may facilitate enzymatic manipulation by inducing structural changes in normal DNA.

The DNA sequence of the 2nd attTn7 site from S. marcescens of plasmid pSIQ 1::Tn7 did not show any homology with either left or right ends of attTn7 sites from other bacteria. A homology search with the EMBO DNA sequence data bank revealed no strong DNA homology with any known DNA sequences. However, a less convincing homology of 25/30 bp with the MetG tyrosyl tRNA synthetase of E. coli was found (see figure 6.6). Since the right end junction fragment from A. tumefaciens was not obtained it is not possible to make any comparison with the attTn7 sites of E. coli, K. pneumoniae and S. marcescens (1st site). Therefore it may be of interest to obtain the DNA sequence of the right end of attTn7 from A. tumefaciens for this comparison.

CHAPTER 7

Effect of Altered Interdomain Sequence and the Deletion of glmS Transcriptional Terminator in the attTn7 on the Transposition of Tn7

7.1 Introduction:

The purpose of the experiments described in this chapter was to understand the interaction between the DNA encoding the carboxy-terminal domain of glmS (glmS box) and the point of insertion of Tn7 in the attTn7 DNA.

The DNA sequence comparison of attTn7 sites from the chromosomes of K. pneumoniae and S. marcescens with the attTn7 from the chromosome of E. coli (Lichtenstein and Brenner, 1982) shows that in all cases the insertion site of Tn7 is located at a considerable distance from the 36 bp conserved sequence encoding the carboxy-terminus of glmS. In the three cases, the point of insertion of Tn7 (the 5 bp target sequence) was found either 2 or 3 integral helical turns of the DNA from the last nucleotide of the conserved domain of 36 bp glmS. This situation suggests some analogy to the well documented cases of gene regulation by two operators separated by an integral number of helical turns (Hochschild and Ptashne, 1986; Irani et al., 1983; Ptashne, 1986; Takahashi et al., 1985). In such cases two proteins

are shown to bind with two conserved DNA sequence elements, facing on the same side of the DNA double helix. The protein-protein interaction allows the intervening DNA sequence to loop out and this brings the two functional domains of a gene together for gene regulation at a distance (Ptashne, 1986).

In the case of attTn7, whilst there is no strong conserved domain at the point of insertion of Tn7, a 5 bp sequence (5'-GCNGG-3') motif located at the insertion point of Tn7 in the attTn7 sites of E. coli and K. pneumoniae was also found 3 bp downstream in the attTn7 of S. marcescens. From the data it was reasonable to suggest that there may be a possible interaction between two domains in the attTn7 site which involve a looping out mechanism. To test the requirement for an integral number of helical turns between the 36 bp conserved sequence and the actual insertion site, it was necessary to separate these two domains by a variable distance. It was then planned to test the effect of these variable spacings on Tn7 transposition using the M13 transduction frequency assay.

Another objective was to delete the DNA sequence at the insertion point of Tn7 in the 70 bp attTn7 DNA sequence; these could not be deleted by exonuclease III treatment from the left (phoS) end of the primary attTn7 (see chapter 4, section 4.1) and thus the requirement for sequences present at the actual insertion site had not been demonstrated. In the following sections of this chapter,

strategies are discussed to obtain the desired target clones which were then tested using the M13 transduction frequency assay as described in chapter 3.

7.2 Generation of ClaI RE Site 5 bp Downstream of the 36 bp glmS Box in 70 bp attTn7 Site (MAS):

To facilitate the construction of altered interdomain spacing in the 70 bp attTn7 DNA (MAS), a mismatched 17-mer oligonucleotide was used to introduce a specific change at a specified position to alter nucleotide C into nucleotide G. The alteration of this nucleotide in the (+) strand of M13 phage vector would have allowed to generate a ClaI RE site at this position (see figure 7.1).

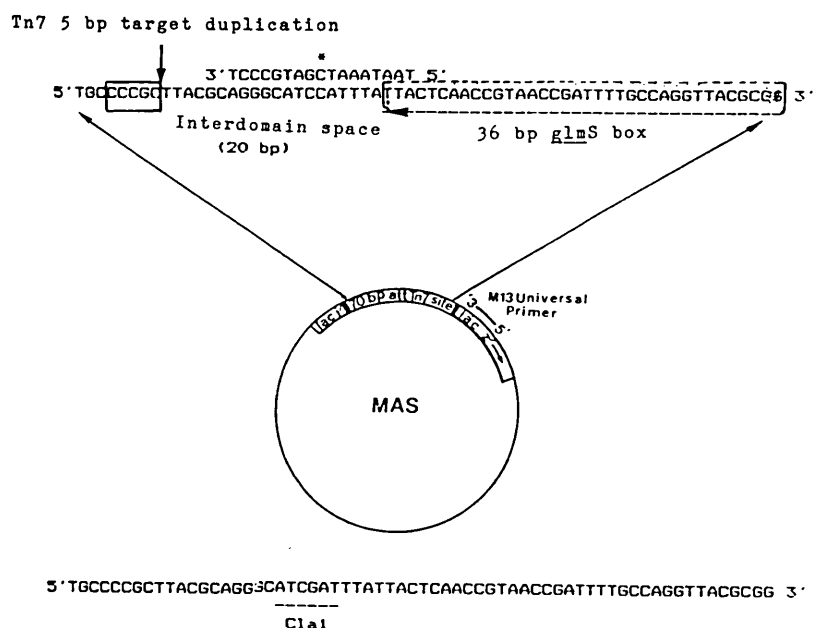
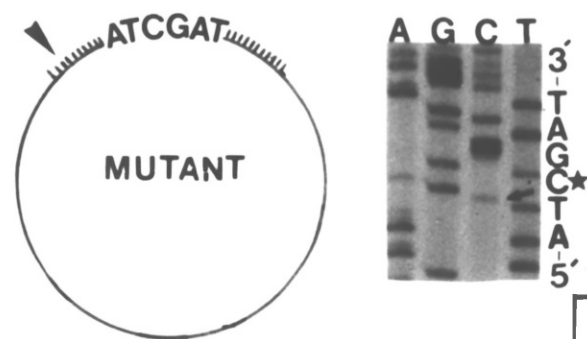
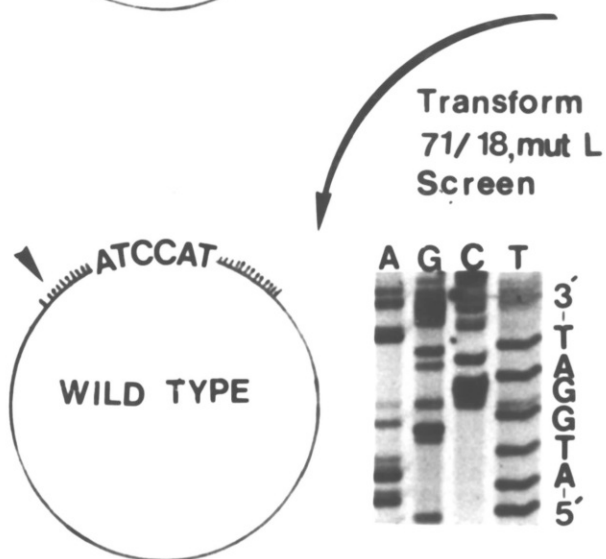
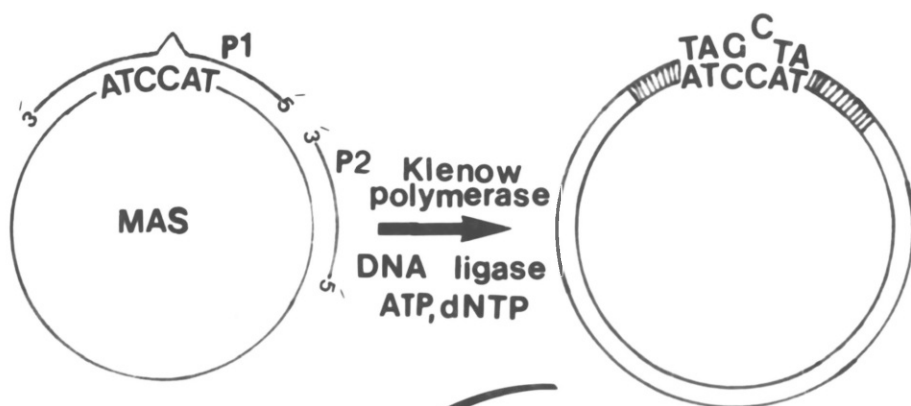


Figure 7.1 Sequence of the mutagenic primer above the mini attTn7 site sequence. The target area for mutagenesis is denoted by an asterisk. The mismatch will be incorporated into RF after complementary strand synthesis. The single base substitution from C to G in the plus strand of MAS will generate a ClaI RE site in the interdomain sequence of MAS.

Figure 7.2 Strategy and results for introducing the ClaI RE site inbetween the end of the glmS coding region and the point of insertion of Tn7. The figure shows the G to C transition in the synthesised strand of MAS. Therefore, this change corresponds to the alteration of C to G in the plus strand. The relevant part of the sequencing gel for wild-type and mutated derivative is shown.



KEY

► Tn7 insertion point

▤ mini *att* Tn7 site

P1 mismatched heptadecamer

P2 M13 universal primer

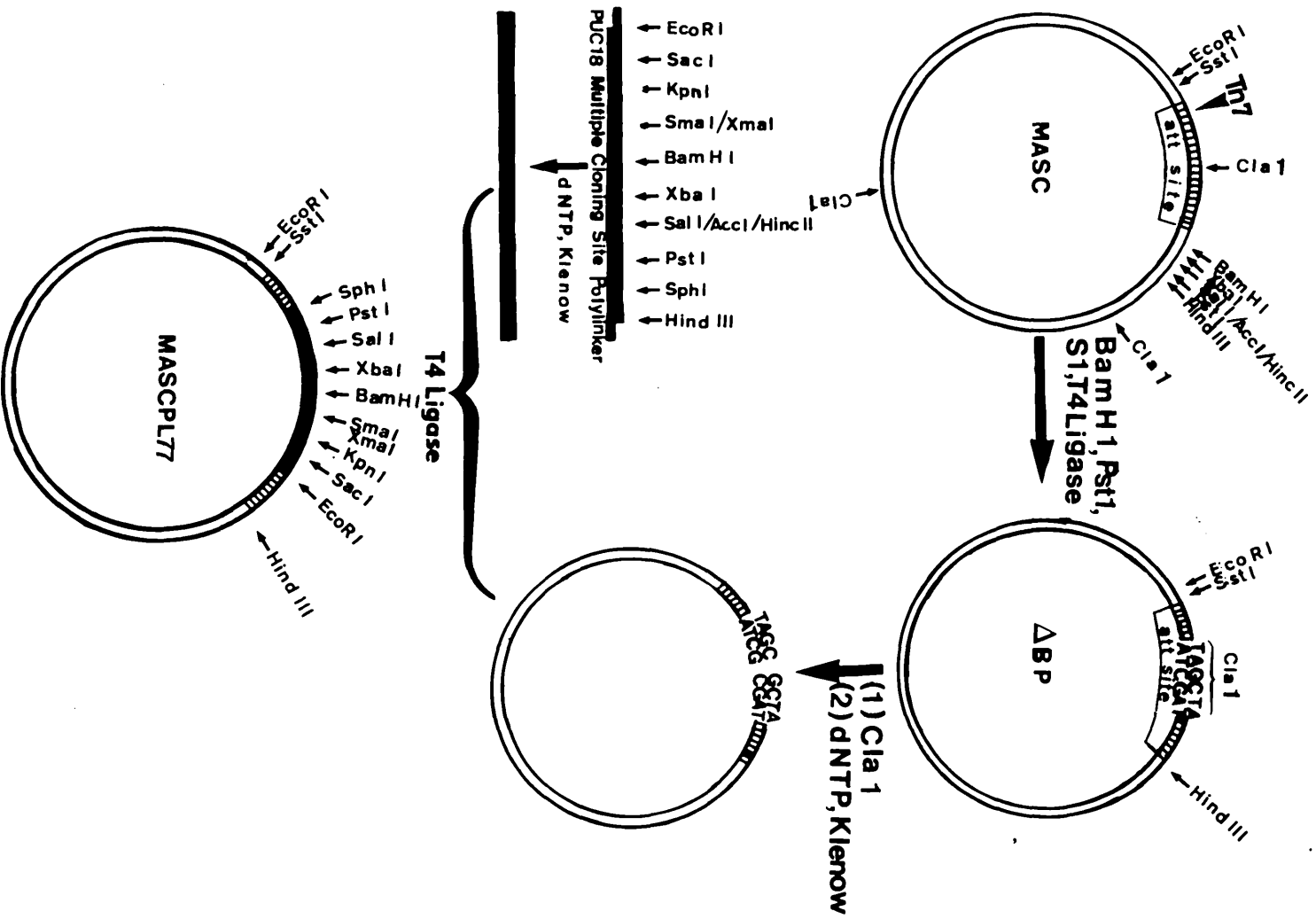
The mutant clone was isolated from the wild-type by using [³²P]-dCTP labelled oligonucleotide as a hybridisation probe (chapter 2, section 2.2.3A). Seven positive clones were identified and were subjected to ClaI RE digestion. The mutant clone named MASC had acquired the ClaI RE site at the expected position, thus yielding a novel 600 bp fragment upon cleavage with ClaI RE. The C to G conversion was confirmed by DNA sequence analysis of MASC and figure 7.2 shows the relevant portion of the sequencing gel.

M13 phages containing the altered nucleotide in the attTn7 sequence (MASC) were tested for transposition of Tn7 using the transduction assay and the frequency was found to be similar to that of MAS. MASC (RF) DNA was used to construct a set of desired clones, the strategies of which are discussed below.

7.3 Construction of Altered Spacing Clones Between the End of 36 bp glmS Box and the Point of Insertion of Tn7 in MASC:

A 55 bp pUC18 multiple cloning site (HindIII RE-EcoRI RE) polylinker was employed for cloning in the newly generated ClaI RE site in the MASC. The strategy for cloning the polylinker at the ClaI RE site is shown diagrammatically in figure 7.3.

Figure 7.3 Strategy for cloning a 55 bp polylinker at the newly generated ClaI RE site in mini attTn7 DNA. The pUC18 multiple cloning site polylinker was employed for cloning purposes. ΔBP was partially cleaved with ClaI RE and the DNA isolated from a 0.8% agarose gel for cloning purposes. MASCP~~L~~77 contains the 77 bp distance between the end of the sequence encoding the C-terminal domain of glmS and the point of insertion of Tn7.



This polylinker had 11 RE sites of which 7 (PstI, SalI, AccI, HindII, XbaI, BamHI and SstI) were already present in the polylinker of M13, mp10 derived clone MASC. To destroy these RE sites, a 20 bp BamHI RE and PstI RE fragment were removed from the MASC and this clone was named ΔBP. The presence of two ClaI RE sites in M13 phage vector, mp10 made the task for cloning in the attTn7 site difficult. Therefore ΔBP (RF) DNA was partially cleaved with ClaI RE and the linearised 7.3 Kb fragment was isolated from an agarose gel for cloning purposes.

The 2 bp (5'GC 3') overhanging sequence of the ClaI RE sites in clone ΔBP was filled in using DNA Polymerase, 1-Klenow fragment; this sequence was thus duplicated at the ClaI RE ends of ΔBP. The 55 bp pUC18 multiple cloning site polylinker (HindIII-EcoRI) was blunt ended in the same way and cloned into the blunt ended ClaI RE site of ΔBP. The transformants were screened using the [³²P]-dCTP labelled HindIII-EcoRI polylinker as a hybridisation probe. Three positive clones were analysed by DNA sequencing and in all cases the insert was found to be in one orientation at the expected position. The obtained clone was named MASCPL77. In this clone the Tn7 insertion site is placed 77 bp in total away from the translational stop codon of the glmS. The 2 bp duplication at the cloning site in ΔBP generated an EcoR1 RE site at the junction of the insert and the attTn7 site DNA in MASCPL77. Two other RE sites HindIII and SstI also occurred twice in the clone

MASCPL77.

Four recombinant clones MASCPL73A, MASCPL73B, MASCPL69 and MASCPL32 with 73, 73, 69 and 32 bp respectively of variable distance between the insertion point of Tn7 and the reference point of the conserved domain (36 bp glmS box) were obtained by using the MASCPL77(RF) DNA. MASCPL43 was obtained by experiments described in section 7.5.

The construction of these MASCPL77 derivatives was achieved by digesting the MASCPL77(RF) DNA with KpnI RE for MASCPL73A, BamHI RE for MASCPL73B, BamHI RE for MASCPL69 and KpnI RE plus SphI RE for MASCPL 32. After restriction endonuclease digestion, the DNA was treated with S1 nuclease to remove the single-stranded overhanging ends and then ligated using T4 DNA ligase. The DNA was transformed to competent bacterial cells of TG2 as described in chapter 2, section 2.2.1(i).

M13 plaques were picked from each plate and single-stranded DNA prepared for DNA sequence analysis (ddTTP reaction only) according to the Sanger dideoxy chain termination method. Figure 7.4 shows the autoradiograph of the DNA sequencing gels of obtained clones. Figure 7.6 shows the DNA sequence of each clone. In clone MASCPL69 an additional 4 bp were digested by S1 nuclease. All clones identified by DNA sequencing were then subjected to the M13 transduction frequency assay and results are presented in table 7.1 (A).

Figure 7.4 DNA sequencing gels of altered spacing clones (MASCPL77 and derivatives). The regions between arrows show the sequence of insert (pUC18 polylinker) at ClaI RE site of MASC. The sequence shown between stars in autoradiograph F is a 12 bp in vivo insert (see figure 7.6 and section 7.5 for details).

A: MASCPL77
B: MASCPL73A
C: MASCPL73B
D: MASCPL69
E: MASCPL32
F: MASCPL43

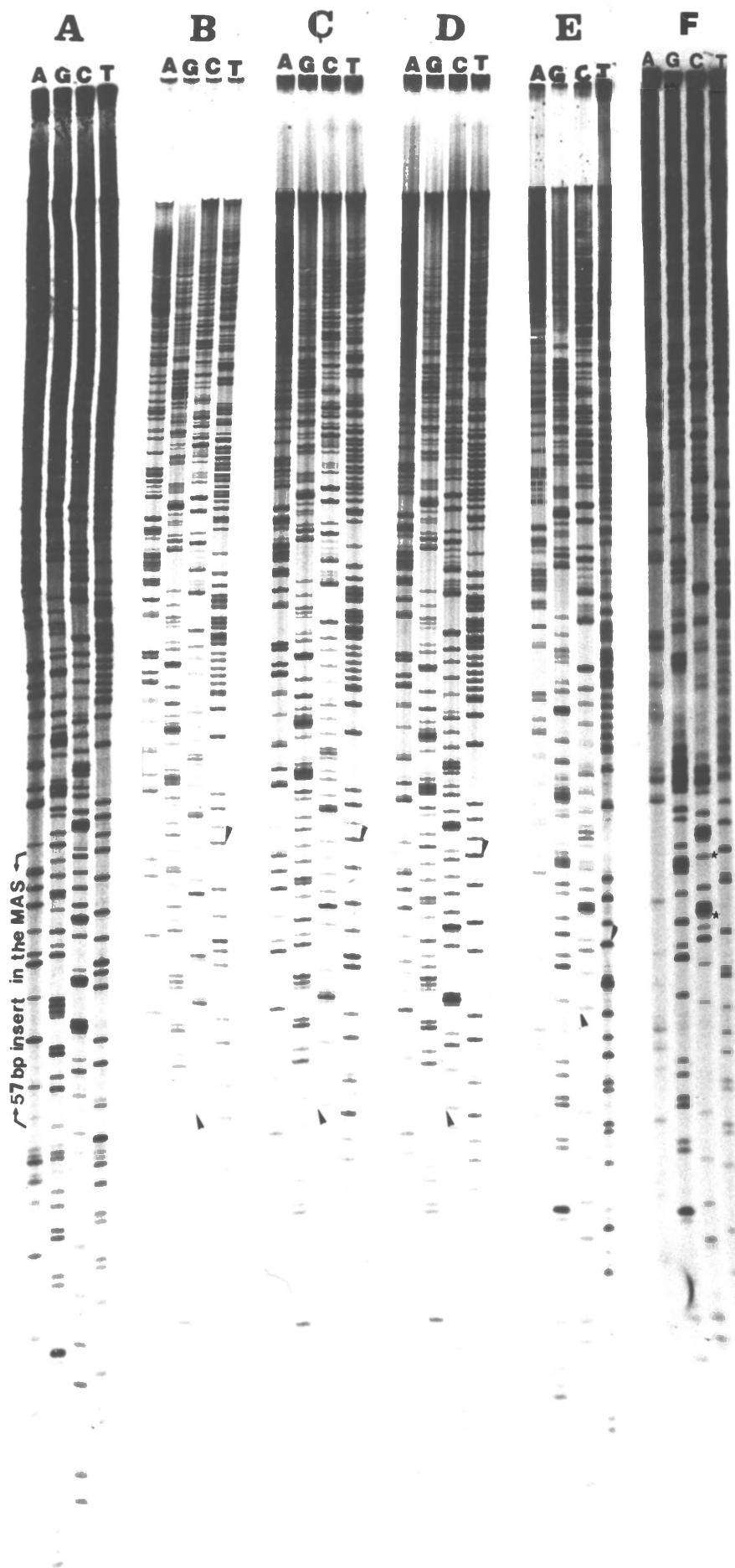
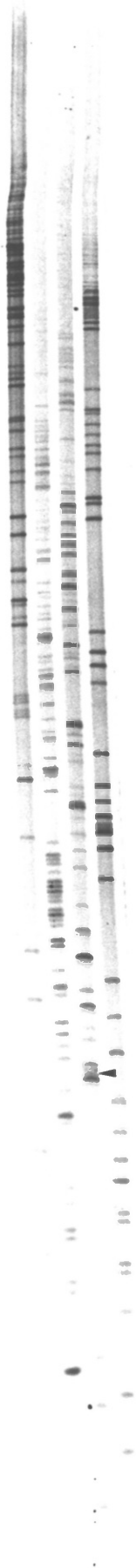


Figure 7.5 DNA sequencing gels of deletion mutants $\Delta 1\text{LOOP}$, $\Delta 2\text{LOOP}$ and $\Delta 3\text{LOOP}$. Arrows indicate the junction of attTn7 sequences and the adjacent sequences next to deletion point in each clone. The DNA sequences below the arrows are of attTn7 sites.

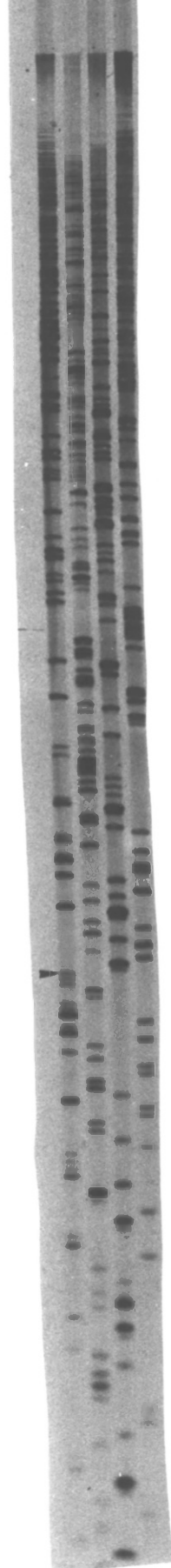
$\Delta 1$ LOOP

A G C T



$\Delta 2$ LOOP

A G C T



$\Delta 3$ LOOP

A G C T



Figure 7.6 DNA sequence of altered spacing clones (MASCPL77 and derivatives) and three deletion clones ($\Delta 1$ LOOP, $\Delta 2$ LOOP and $\Delta 3$ LOOP). The deletion clones were obtained in experiments described in section 7.4. Altered spacing clones (as shown in parenthesis) contain various lengthed pUC18 multiple cloning sites polylinker sequence (bold typing) between the end of the glmS coding region and the point of insertion of Tn7 in the attTn7. MASCPL43 is MASCPL32 derivative and contains an in vivo insert of 12 bp (shown in a box) and 1 bp deletion. The deletion clones $\Delta 1$ LOOP and $\Delta 2$ LOOP contain the DNA sequence of the plasmid pGC2 (GC clamp) and M13 polylinker respectively (also shown in bold typing) which is placed in proximity to the glmS sequence and no potential cruciform structure exist in these clones (see text for detail). $\Delta 3$ LOOP contains the DNA sequence of pUC18 multiple cloning sites plus M13 polylinker and contains a novel potential cruciform structure in proximity to the glmS sequence in the 70 bp attTn7 site.

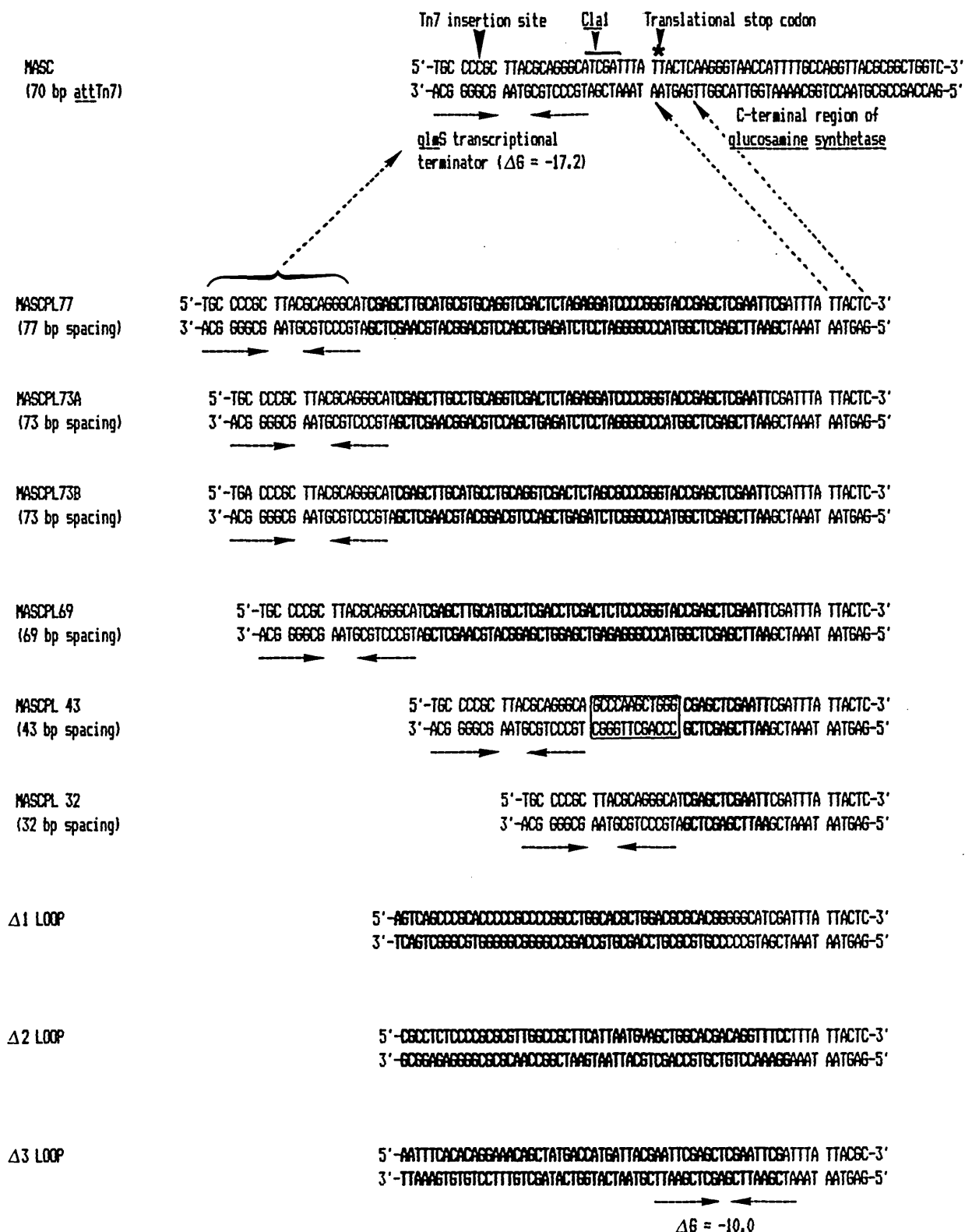


Table 7.1 Results of the M13 transduction frequency assays on altered spacing (MASCPL77 and derivatives) and deletion clones ($\Delta 1$ LOOP, $\Delta 2$ LOOP and $\Delta 3$ LOOP).

A: The assays were performed using 71/18 (ColE1::Tn7) as Tn7 donor strain and TG2 as recipient strain for scoring trimethoprim transduced colonies.

Target	Independent trials			
	1	2	3	4
MAS	1.1E-4	5.0E-5	1.3E-4	5.0E-5
MASC	1.0E-4	3.6E-5	1.1E-4	5.0E-5
MASCPL77	6.3E-5	1.5E-5	2.3E-4	5.0E-5
MASCPL73A	7.5E-5	3.1E-5	5.1E-4	3.5E-5
MASCPL73B	1.1E-4	3.5E-5	2.8E-4	2.1E-5
MASCPL69	1.9E-4	5.6E-5	5.6E-5	1.1E-4
MASCPL43	1.5E-4	4.1E-5	1.7E-4	5.0E-5
MASCPL32	2.3E-4	5.0E-5	8.5E-5	1.5E-4
$\Delta 1$ LOOP*	7.5E-6	1.1E-6	5.0E-6	8.0E-7
$\Delta 2$ LOOP*	2.5E-6	1.4E-6	8.0E-7	1.0E-7
$\Delta 3$ LOOP*	8.0E-5	6.2E-5	1.3E-4	5.3E-5

B: Relative values of transduction frequencies of altered spacing and deletion clones to that of control (MAS).

Target	Independent trials				Mean relative value
	1	2	3	4	
MAS	1.0E 0	1.0E 0	1.0E 0	1.0E 0	1.0E 0
MASC	9.0E-1	7.2E-1	8.4E-1	1.0E 0	8.6E-1
MASCPL77	5.7E-1	3.0E-1	1.7E 0	1.0E 0	8.9E-1
MASCPL73A	6.8E-1	6.2E-1	3.9E 0	7.0E-1	1.4E 0
MASCPL73B	1.0E 0	7.0E-1	2.1E 0	4.2E-1	1.0E 0
MASCPL69	1.7E 0	1.1E 0	4.3E-1	2.2E 0	1.3E 0
MASCPL43	1.3E 0	8.2E-1	1.3E 0	1.0E 0	1.1E 0
MASCPL32	2.0E 0	1.0E 0	6.5E-1	3.0E 0	1.6E 0
$\Delta 1$ LOOP	6.8E-2	2.2E-2	3.8E-2	1.6E-2	3.6E-2
$\Delta 2$ LOOP	2.2E-2	2.8E-2	6.1E-3	2.0E-3	1.4E-2
$\Delta 3$ LOOP	7.2E-1	1.2E 0	1.0E 0	1.0E 0	9.8E-1

* Clones $\Delta 1$ LOOP, $\Delta 2$ LOOP and $\Delta 3$ LOOP were obtained in those experiments described in section 7.4.

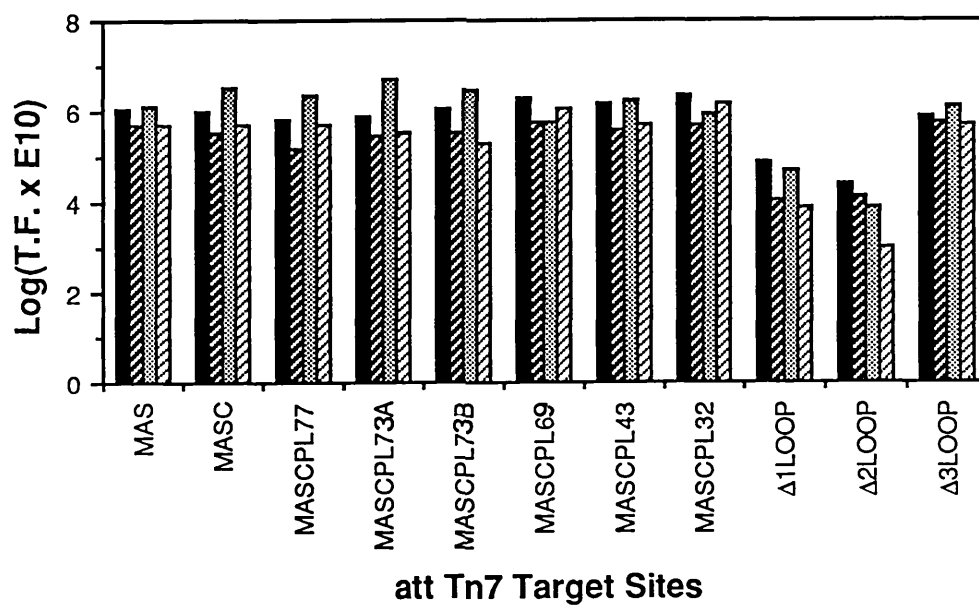
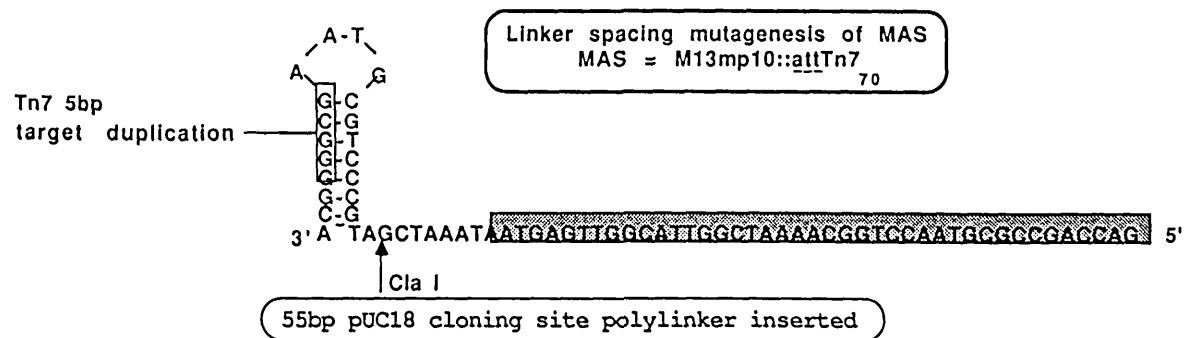


Figure 7.7 Graphical representation of the results of M13 transduction frequency assays on altered spacing and deletion clones. This graph was plotted with the same procedure as described in figure 3.2.



Number of helical turns	Potential cruciform	Transduction frequency
7.3	Yes	w.t.
7.1	Yes	w.t.
7.1	Yes	w.t.
6.5	Yes	w.t.
4.0	Yes	w.t.
3.0	Yes	w.t.

Figure 7.8 Results of linker spacing mutagenesis of MAS.

Table 7.1 (B) shows the relative and mean relative values of transduction frequency of each clone to the wild-type (MAS). The transduction frequency results of these altered spacing and deletion clones are presented in graphical form in figure 7.7. Results of altered spacing experiments are summarised in figure 7.8 and table 7.2.

The results clearly show that in recombinant clones (MASCPL77, MASCL73A, MASCPL73B, MASCPL69 and MASCPL32) there is no apparent effect of variation of interdomainal spacing on the transduction frequency and hence Tn7 transposition frequency.

Table 7.2 Effect of (i) altered spacing, (ii) potential cruciform structure on the transposition frequency of Tn7 as established by using the M13 transduction frequency assay.

Name of clone	Number of helical turns between the end of the 36 bp <u>glmS</u> box and the beginning of target sequence	Potential cruciform structure	Transduction frequency relative to wild-type (MAS)
MAS	2.0	yes	1.00
MASC	2.0	yes	0.86
MASCPL77	7.3	yes	0.89
MASCPL73A	7.1	yes	1.47
MASCPL73B	7.1	yes	1.05
MASCPL69	6.5	yes	1.35
MASCPL43 [†]	4.0	yes	1.10
MASCPL32	3.0	yes	1.67
Δ1LOOP	-	no	0.036
Δ2LOOP	-	no	0.014
Δ3LOOP	-	yes	0.980

[†]MASCPL43 was obtained in experiments described in section 7.5.

Strategies were now adopted to delete the actual Tn7 insertion site, i.e. glmS transcriptional terminator in the attTn7 and test the effect of these deletions on the transduction frequency. This was designed to define any functional role of the 3' non-coding region in the attTn7.

7.4 Construction of Deletion Mutants in the glmS Transcriptional Terminator in 70 bp attTn7 Site and Effect of These Deletions on Transduction Frequencies:

The results of experiments described in section 7.3 show that an integral helical turn spacing between the 36 bp glmS box and the point of insertion of Tn7 in attTn7 is not required for the high frequency transposition of Tn7. The experiments described in this section address the possibility of whether the Tn7 insertion site in attTn7 was required in conjunction with the 36 bp glmS box. The strategies adopted to define the role of this 3' non-coding sequence in the attTn7 were deletions of various magnitudes from the right end of the attTn7.

Three deletion mutants $\Delta 1\text{LOOP}$, $\Delta 2\text{LOOP}$ and $\Delta 3\text{LOOP}$ having deletions of 15, 24 and 20 bp from the right end of 70 bp attTn7 are described in this section for clarity and convenience. However, the deletion mutant $\Delta 1\text{LOOP}$ was not obtained in the experiments described in this section. In contrast the $\Delta 1\text{LOOP}$ was obtained by chemical treatment on 70 bp attTn7 DNA, the procedure for which is discussed in

chapter 8, section 8.2.

Δ 2LOOP was obtained by digesting the MASC (RF) DNA with ClaI RE and SstI RE. The linearised DNA was then treated with S1 nuclease to remove the single-stranded overhanging DNA. The DNA was ligated with T4 DNA ligase and transformed to bacterial strain TG2 as described in chapter 2, section 2.2.1(i). The transformants were screened by DNA sequencing and figure 7.5 shows the autoradiograph of the obtained deletion mutant. Figure 7.6 shows the DNA sequence of this clone as well as that of Δ 1LOOP. In Δ 2LOOP a linked deletion of approximately 200 bp in the lacI region of M13, mp10 was also found. The reason for this deletion is unclear.

Δ 1LOOP and Δ 2LOOP were tested using the M13 transduction frequency assay and the results are shown in table 7.1 (A). Table 7.1 (B) shows the relative and mean relative transduction frequency values for these deletion mutant clones to the wild-type (MAS).

The results show a drop of approximately 27 and 71 fold in the transduction frequency value of Δ 1LOOP and Δ 2LOOP, respectively, in comparison to the wild-type transduction frequency of MAS. It seems that the DNA sequence in the 3' non-coding region of the 70 bp attTn7 is also responsible for controlling the high frequency transposition of Tn7. Another possible factor for this drop may be the involvement of the glmS transcriptional terminator in the

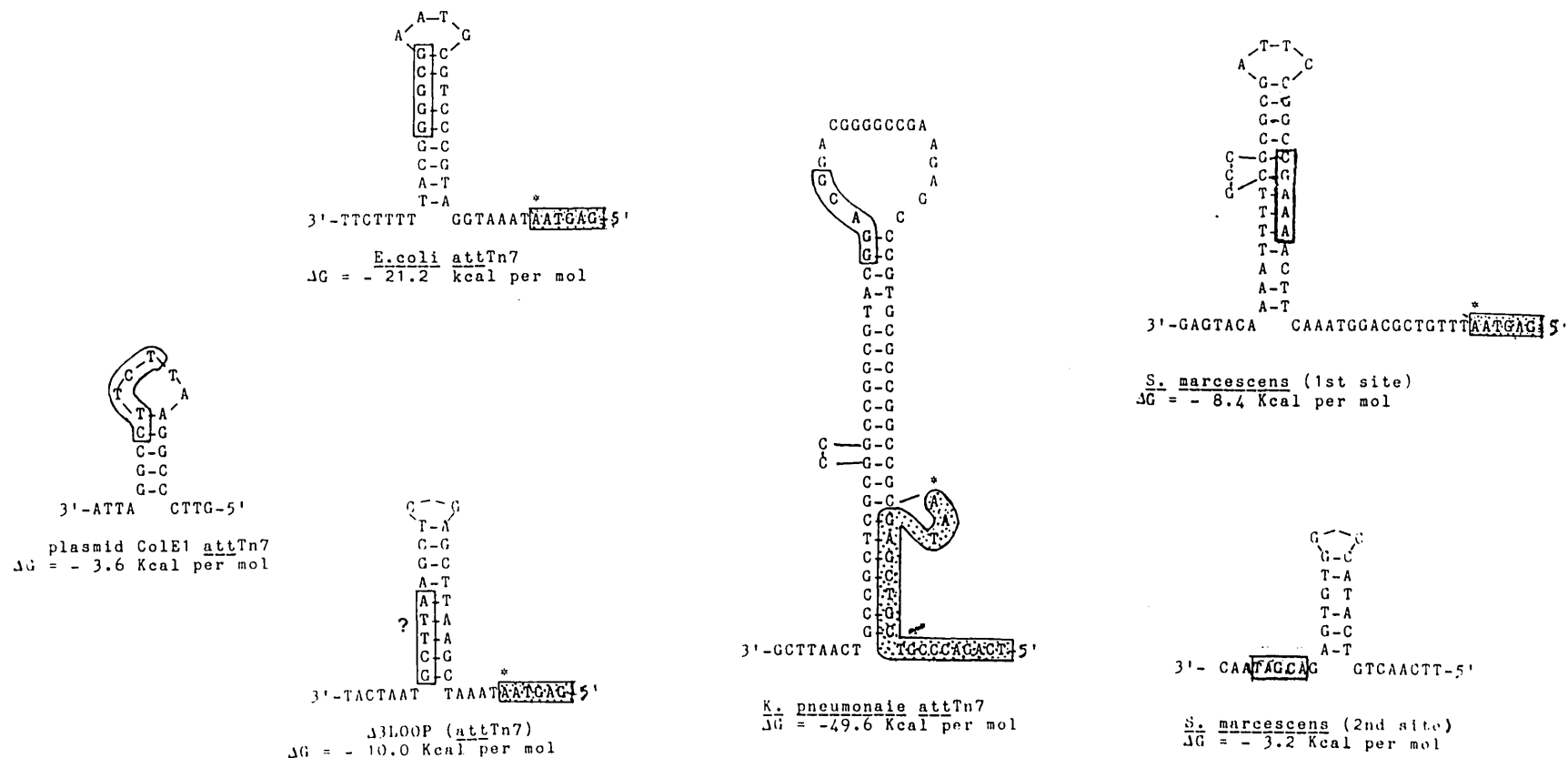


Figure 7.9 Potential cruciform structures around the 5 bp target duplication in attTn7 sites. The DNA sequence at insertion site of *E. coli* and plasmid ColE1 is taken from Lichtenstein and Brenner (1982) and Walker et al., 1984. The 5 bp target sequence is shown in a box and in $\Delta 3L00P$ this sequence is not known. The DNA sequence encoding the carboxy-terminus of *glmS* is shown in shaded boxes and star is the translational stop codon of *glmS*.

attTn7 site, as the deletion in both cases has removed the strong potential cruciform structure with a thermodynamic value of $\Delta G = -19.2$ Kcal per mol.

In order to pursue this further, a deletion mutant clone $\Delta 3\text{LOOP}$ was constructed which has a deletion of 20 bp from the right end of 70 bp attTn7. This clone was obtained by digesting the MASCPL77(RF) DNA with SstI RE. The linearised DNA was then ligated in vitro. By doing this it was possible to generate an EcoRI RE site in the attTn7; thus the sequence obtained has the ability to form a cruciform structure ($\Delta G = -10.0$ Kcal per mol) with the DNA of the M13, mp10 (see figure 7.6). After ligation the transformants were screened by DNA sequence analysis (ddTTP track only) and the desired clone, was identified.

The deletion mutant clone $\Delta 3\text{LOOP}$ was tested using the M13 transduction frequency assay and the consequent results show a wild-type frequency (see table 7.1). These results indicate that there is a structure, not a sequence, in the 3' non-coding region of the attTn7 involved in governing the high frequency transposition of Tn7 in addition to the 36 bp glmS box.

The transduction frequency results obtained on the varying spacing clones (MASCPL77 and derivatives) and the deletion clones ($\Delta 1\text{LOOP}$, $\Delta 2\text{LOOP}$ and $\Delta 3\text{LOOP}$) were consistent. These can be summarised with respect to each other as shown in table 7.2. Figure 7.7 shows the graphical representation of

the results of M13 transduction frequency assays on these deletion clones.

From the analysis of these deletion mutants spanning the point of Tn7 insertion in E. coli attTn7, it is evident that the actual point of insertion is not required for high frequency transposition. However data do suggest that in addition to the 36 bp glmS box defined previously there may be a general requirement for an inverted repeat (potential cruciform) in the vicinity of insertion sites. It is noteworthy that in the attTn7 sites of K. pneumoniae and S. marcescens (both sites) the point of insertion is also near to an inverted repeat (see figure 7.9). $\Delta 1\text{LOOP}$ and $\Delta 2\text{LOOP}$ lacking an inverted repeat have reduced transduction frequency 27 and 71 fold respectively, in comparison to $\Delta 3\text{LOOP}$ which has such an inverted repeat. The DNA sequence of $\Delta 3\text{LOOP}$ shows no homology to MAS at the expected Tn7 insertion point (i.e. 20 bp away from the 36 bp glmS box).

There seems to be two domains in the attTn7 which probably participate in a sequence-specific and structure-specific manner during Tn7 transposition. It may be that the sequence-specific domain (36 bp glmS box) confers the specificity for Tn7 transposase and the structure-specific domain apparently serves to present the target DNA as an appropriate point for the insertion reaction and seems not to be involved in specificity and recognition.

7.5 Where is Tn7 Inserted in Altered Spacing and Deletion Clones?:

To further understand the interaction between the 36 bp glmS box and the point of insertion of Tn7, it is necessary to investigate the point of insertion of Tn7 in the altered spacing and deletion mutants of attTn7. Since M13::attTn7^r phages transducing Tp^r had deletions it was very difficult to obtain the M13 clones with intact Tn7-attTn7 junctions. However, attempts were still made to purify these phage particles. It was decided to streak purify the Tp^r transduced colonies on MH plates containing trimethoprim. Bacterial colonies were repeatedly streaked alternately for colonies and plaque formation. In this selection 60% of the Tp^r colonies were found to be unable to produce M13 plaques. This can be explained in two ways, either, in these colonies intact Tn7 has transposed from M13::attTn7::Tn7 to the chromosome or the phage particles had lost the ability to package the genome.

After repeated experiments the M13 particles transducing trimethoprim were obtained and the DNA sequence analysed (Sanger et al., 1977). Only 3 out of 150 clones analysed were found to contain the left end of Tn7. However, these were not altered spacing clones as wished but did have Tn7 inserted into the expected point of Tn7 insertion (i.e. 20 bp away from 36 bp glmS box, data not shown). This suggested that mini attTn7 can provide a site-specific target for Tn7 transposition when cloned in M13 phage

vectors.

A clone MASCP43 was found among the 12 MASCP32::T^rp clones analysed by DNA sequencing (see figure 7.6). This clone has acquired an in vivo insert of 12 bp and a deletion of 1 bp in MASCP32. In MASCP43 the wild-type point of insertion of Tn7 was 43 bp away from the 36 bp glmS box. This in vivo rearrangement within MASCP32 seems to be very precise, i.e. the 12 bp insert is found 19 bp away from the end of glmS box in MASCP32. The sequence of the original glmS transcriptional terminator is intact in this clone. Therefore, this structure could be explained with the assumption that insertion at this point has led to a near precise excision of Tn7 and leaving a 12 bp insert. No detectable homology of this insert was found with the ends of Tn7. This clone was tested as a target for Tn7 transposition by using the M13 transduction assay and the results are shown in table 7.1.

By following the strategy discussed above it was difficult to obtain the desired clones due to concomitant deletions. However, the cloning of different attTn7 sites in wide host range plasmids (e.g. pTJS75Kn or pRK404) and hopping Tn7 into these attTn7 sites may allow suitable junction fragments for DNA sequencing to be found. Unsuccessful attempts were made to pursue this strategy.

7.6 Conclusions from Altered Spacing and Deletion Experiments:

The results of experiments performed on different altered spacing and deletion clones can be rationalised as follows.

All spacing clones described in section 7.3 failed to provide evidence for any significant role of a specific DNA sequence, an integral of helical turns from the 3' end of the 36 bp glmS box for high frequency transposition of Tn7. The question remains why in E. coli, S. marcescens and K. pneumoniae the point of insertion is located always 2 or 3 helical turns of DNA away from the highly conserved glmS box. It may be that target site selection involves a sort of "ruler" that measures an integral number of helical turns to the beginning of 5 bp target site duplication. Also unknown is the reason of point of insertion of Tn7 at 2 helical turns in E. coli and S. marcescens (1st site) and 3 helical turns in K. pneumoniae.

Whether this difference in the measurement of distance is accompanied by different host factors or if it is due to the nature of DNA sequence remains unclear. In K. pneumoniae attTn7 the 31 bp (i.e. 3 helical turns of DNA) of non-coding region contains 26 GC base pairs. This GC richness may be the requisite of an additional helical turn spacing at point of insertion in K. pneumoniae attTn7 site. This

may be related to the capability of the DNA to form some altered configurations (i.e. Z configuration). For example, the atomic resolution of single crystal X-ray of a GC DNA hexamer [d(CpGpCpGpGpCp)G] has been shown to form left handed Z-configuration (Wang et al., 1979).

The physical analysis of K. pneumoniae attTn7 is necessary to determine such configurations which may in turn determine this spacing. The question regarding the involvement of any specific host factor for variable spacing can be answered in the first stage by testing the virgin attTn7 site from K. pneumoniae in altered hosts, i.e. E. coli or S. marcescens. If a host specific factor is responsible for this phenomenon, then Tn7 insertion would be expected to occur after a distance of 2 helical turns away from the 36 bp glmS box in the absence of this factor in an altered bacterial host. Later, this factor may be purified from K. pneumoniae and can be provided in an expression vector for restoring the Tn7 insertion at 3 helical turns distance in E. coli or S. marcescens. However conclusive evidence for the existence of such a ruler must however await an analysis of the insertions of Tn7 in altered spacing clones and deletion mutants, in particular $\Delta 3L00P$.

Although deletion of the actual site of insertion does not affect the transposition frequency, providing a secondary structure is available within the region, it seems that the 36 bp glmS box can function on its own. However the 36 bp

glmS box has not been fully defined, as the deletions within this box from the left end were not achieved.

Another conclusion drawn from this analysis is that mini 70 bp attTn7 can provide a high frequency and site-specific target for Tn7 transposition of Tn7 when cloned into M13 phage vector.

To extend the understanding of interaction between different domains in the attTn7, it was required to study the contact points of proteins (Tn7 or host-encoded) at the attTn7 site, as manifested by the effect of base substitution on the reactivity of Tn7. In the absence of an established in vitro transposition system, random mutants were generated in the 70 bp attTn7 by in vitro mutagenesis and their effect on the in vivo transduction assay was measured.

It was expected that mutants showing reduced transposition frequency would be clustered in regions which are involved directly in the transposition reaction. These experiments are discussed in chapter 8.

CHAPTER 8

Effect of Nucleotide Substitution in the attTn7 Sequence on Tn7 Transposition and Studies on DNA-Protein Interactions Involving Mini attTn7 Site

8.1 Introduction:

The involvement of particular stretches of DNA in the attTn7 DNA for high frequency transposition of Tn7 were elucidated by the experiments described in chapters 4, 5, 6 and 7. The deletion analysis of the attTn7 site has described the limits of the DNA sequence (Δ R25 and Δ L849) within which domains were responsible for high frequency transposition of Tn7. A detailed analysis of DNA sequence defined by these deletion studies was now required to understand which specific pairs of nucleotides participate in target site selection.

To understand the precise function of DNA sequence in genetic control and regulation, two strategies are commonly employed which are the subject of this chapter. The first method involves a random or specific mutagenesis within short stretches of DNA. Specific amino acid changes can also be made within the DNA binding domains of proteins by altering the specific nucleotides in the protein coding region (Hill et al., 1986; Myers et al., 1985; Smith,

1984). This allows a change in the binding affinities of regulatory proteins with specific domains of DNA, thus altering the spatial configuration of the DNA-protein complex.

In many biological systems the binding of proteins at minor or major grooves within the specific DNA sequence regulates the specific function, for example transcription, DNA replication and DNA recombination. These DNA-protein interactions have been the foci of in vitro studies in recent years. A second approach can also be taken to understand the cooperative binding of proteins to the regulatory sequences of DNA by gel retardation assays or DNase I foot print experiments (Garner and Revzin, 1981; Strauss and Varsharky, 1984; Wu and Crothers, 1984). To understand the binding of proteins (Tn7 encoded) to the attTn7 site, a gel retardation assay was performed which is the subject of discussion in section 8.6 of this chapter.

The main part of this chapter discusses the strategies adopted to define the regions within the 70 bp attTn7 DNA (which may be the potential transposase binding domain) by altering the individual nucleotides. The aim was to generate a large number of independent point or multiple mutants in the region of attTn7 DNA and to confirm the base substitutions by DNA sequencing (Sanger et al., 1977).

It was hoped that an alteration in specific nucleotides in attTn7 would result in a decrease, or maybe an increase, in

the transposition frequency of Tn7 using the M13 transduction frequency. Thus, the regions of DNA responsible for affecting the wild-type transduction frequency would be identified and might define the binding domains for Tn7 and/or host encoded proteins.

8.2 Experimental Approaches Adopted for in Vitro Random Region Specific Mutagenesis:

Of the varieties of approaches available for mutagenesis there was the possibility to use either an oligonucleotide directed or a random chemical mutagenesis techniques. The former has recently become a widely used and an invaluable means of probing DNA function by introducing very specific base changes in a particular stretch of DNA. The aim was to identify the nucleotides responsible for high frequency transposition in the attTn7 whose positions in the target site were not established; thus the use of oligonucleotide directed mutagenesis techniques was not applicable. It was decided instead to perform random saturation mutagenesis studies on the attTn7 DNA.

Two approaches which were taken, with limited success, to generate mutants in the attTn7 site are discussed below. The first method adopted involved the use of forced misincorporation of a deoxy-ribonucleotide during repair of DNA in vitro synthesis of a gap using the error prone DNA polymerase, AMV reverse transcriptase.

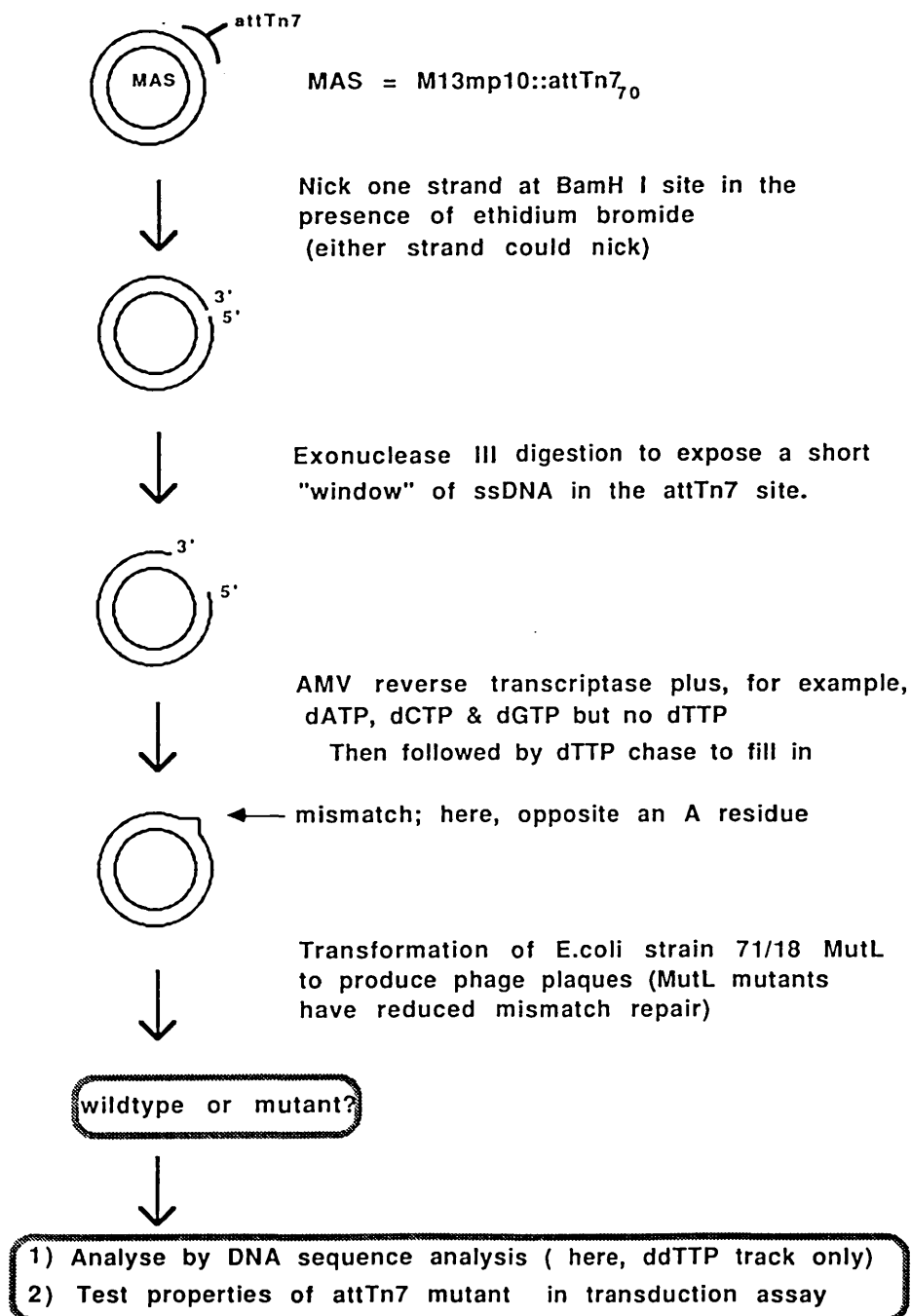


Figure 8.1 Diagrammatic representation of forced misincorporation mutagenesis.

This had previously been used to generate transversion mutants (i.e. GC to CG and GC to TA) in the single-stranded phage ϕ x174 (Zakour and Loeb, 1982). In this method a short gap is generated in vitro in the target double-stranded circular M13 molecules by limited exonuclease III digestion from a site-specific single-stranded nick in the DNA molecules. This plasmid product is then used as a substrate for in vitro misincorporation using an error prone AMV-reverse transcriptase and omitting one or other of the dNTPs. This generates a population of heteroduplex RF molecules with single base pair change which can be used to transform a competent F' bacterial strain to obtain M13 plaques. The details of this procedure are outlined in figure 8.1. This method exploits the following principles:

Step 1: When a supercoiled plasmid or M13 (RF) DNA is digested by a particular restriction enzyme in the presence of ethidium bromide, cleavage takes place first at one or other strand of DNA within the recognition sequence. The CCC molecules are converted into open circular (OC) and are able to take greater amounts of ethidium bromide. This binding blocks cutting of the other strand to yield linear molecules; only half of the molecules will generate a nick at the desired strand. Experiments were performed to optimise the yield of OC molecules as visualised by agarose gel electrophoresis (data not shown). This single-stranded nick can then be used as a substrate for limited Exo III digestion to remove a variable stretch of DNA from

3' to 5' direction. To optimise the yield of the Exo III generated window of single-stranded DNA, different time periods of incubation were followed as discussed in chapter 2, section 2.2.3B(i).

Step 2: The DNA substrate is then used for misincorporation using AMV reverse transcriptase in the absence of one or other of the four dNTPs which is followed by the faithful DNA synthesis in the presence of all four dNTPs.

Step 3: As there is no direct way to select the mutants in the attTn7 DNA, they can only be screened by DNA sequencing reaction (Sanger et al., 1977) or by using the M13 transduction assay as described in chapter 3. Although the final objective was the measurement of transduction frequency of the mutants, it was easier to screen the mutants by DNA sequencing first.

The experiments described in this method were unable to yield the desired results as the DNA sequence reactions of the 55 independent M13 clones picked after mutagenesis had only produced mutant clones in which the substitution was outside the region of interest (data not shown).

The reason for this low frequency yield could be explained by the window selection of target DNA with Exo III treatment. Since the target was only 70 bp, it was difficult to obtain variable-length start points for error prone DNA synthesis within this 70 bp window.

This method was abandoned in favour of the techniques of Myers et al . (1985) for random mutagenesis which involves the physical separation of in vitro chemically mutated DNA from non-mutated DNA, on a denaturing gradient gel in the absence of phenotypic selection. This physical separation allows the enrichment of mutants, thus overcoming the problem of laborious DNA sequence analysis of a large number of clones.

Myers first performed this method on a DNA fragment containing the mouse beta-globin promoter and since then no reports on the applicability of this method on any other DNA fragments have been published. However this method was employed with the hope of obtaining mutants in the 70 bp attTn7 DNA.

The physical separation of mutant fragments in this method exploits the observation that the position in the denaturing gradient gel at which the decrease in mobility is observed corresponds to the melting temperature (T_m) of the DNA fragment. Mutations in the DNA of interest result in an altered T_m , and therefore the mutated fragments migrate with altered mobilities through the denaturing gradient acrylamide gel. The mutants can be separated from the wild-type DNA fragment by first cloning the DNA sequence under study next to a GC rich DNA sequence (GC clamp); this prevents the complete strand dissociation as the DNA fragments (wild-type or mutant) proceed through the denaturing gradient gel.

The first step in this method involves the determination of the melting behaviour of the target DNA fragment attached to the GC clamp by examining the pattern of electrophoresis in a gel containing a gradient of denaturant perpendicular to the direction of electrophoresis. On the right side of the gel where the denaturant concentration is high the DNA fragment melts upon entering the gel and has retarded mobility and an intermediate level of penetration of DNA molecules into the gel occurs at intermediate positions. The resulting pattern of electrophoresis is reminiscent of a hyperchromicity curve. The denaturant concentration at the midpoint of an inflection corresponds to the T_m of a particular melting transition and can be translated into temperature-equivalent. In practice, simply knowing the denaturant concentration at a transition gives all the information that is needed to determine the best concentration of denaturant to use in the preparative parallel denaturing gradient gels. For the attTn7 DNA (cloned only in one orientation in front of the GC clamp, plasmid pGC270) the melting occurred at 60% denaturant gradient while the gel was run at a fixed temperature 60 °C (see figure 8.2).

The steps involved in the procedure of Myers's mutagenesis technique are outlined below.

Step 1: Single-stranded target DNA is mutagenised (MAS in this study) with different chemicals (e.g. formic acid,

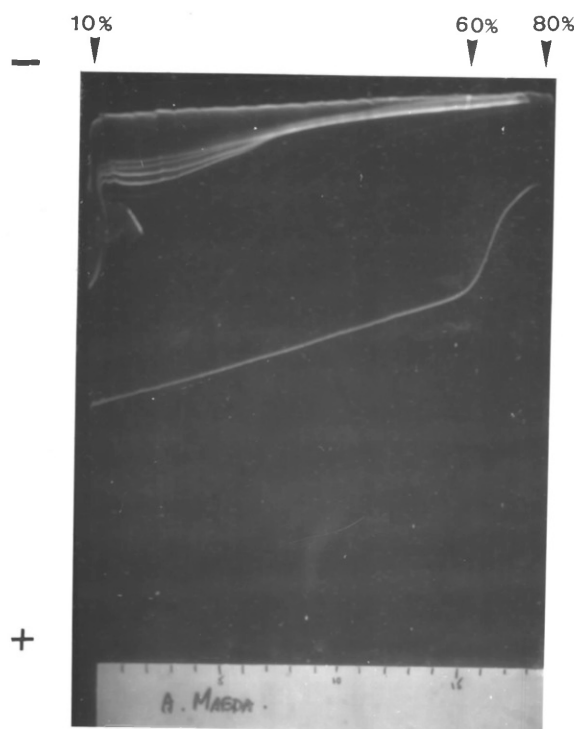


Figure 8.2 A perpendicular gradient (10% to 80%) denaturant acrylamide gel. Approximately 50 μ g HindIII and EcoRI RE digested DNA fragment containing the 70 bp attTn7 plus 250 bp GC clamp was loaded on the gel and run in 60°C oven at constant 150 V for 12 hours in preheated 1X TBE buffer. It can be seen that the melting of the fragment occurs around 60% denaturant gradient.

hydrazine and nitrous acid) for appropriate timings [see chapter 2, section 3.2.3B (ii)].

Step 2: After removal of the chemicals, an oligonucleotide is annealed to the DNA and then extended with AMV-reverse transcriptase using all four dNTPs.

Step 3: The duplex target DNA fragment is then excised with two REs (in this study BamHI and EcoRI) and cloned into the compatible ends in plasmid vectors pGC 1 or pGC 2, (Ap^r) which contain a cloning site polylinker in either orientation with reference to the GC clamp.

The ligated DNA is transformed to competent bacterial cells to select for Ap^r colonies. The pooled DNA from these colonies is digested with two REs to excise the DNA of interest along with the GC clamp which is then electrophoresed on the preparative parallel denaturing gradient gel.

After employing the above strategy, the mutated attTn7 DNA was loaded onto a preparative parallel denaturing gradient gel as described in chapter 2 section 2.2.2C(v). A very faint band appeared above the wild-type fragment (data not shown). The faint DNA band was isolated from the acrylamide gel as described in chapter 2, section 2.2.2B(vii). The DNA fragment was then cloned into M13 phage vector, mp11 at HindIII and EcoRI RE sites for DNA sequence analysis by dideoxy chain termination method (Sanger et al., 1977).

The DNA sequence analysis of 50 clones has produced only three mutant clones, all having a deletion of 15 bp rather than base substitutions in the attTn7 DNA. This deletion was in the region for a potential cruciform structure and therefore removed the 15 bp from the left end of the attTn7 of glmS transcriptional terminator which also included the point of insertion of Tn7 (5'-GCGGG-3'). The mutant (named Δ 1LOOP) is discussed in chapter 7, section 7.3.

There could be several reasons for the poor yield of mutants using this method. First, the conditions involved in the running of the gradient gels were optimised only for the mouse beta-globin promoter region attached to the GC clamp. The melting behaviour of other different DNA fragments may be variable due to the GC contents of the individual target DNAs. Second, due to the lack of a proper apparatus to run the fixed-temperature denaturing gradient gel, the electrophoresis was carried out in an incubator. It was always difficult to maintain the exact conditions described by Myers et al. (1985). Third, this procedure involves too many cloning steps which were time consuming. Fourth, there was no selection for the insert in the plasmid vector containing the GC clamp (i.e. pGC 1 or pGC 2). Fifth, the target DNA was too small to work with as the isolation procedures are laborious. For these reasons, no further attempts were made to optimise the conditions of this procedure and the method abandoned in favour of sodium bisulphite treatment which is discussed below.

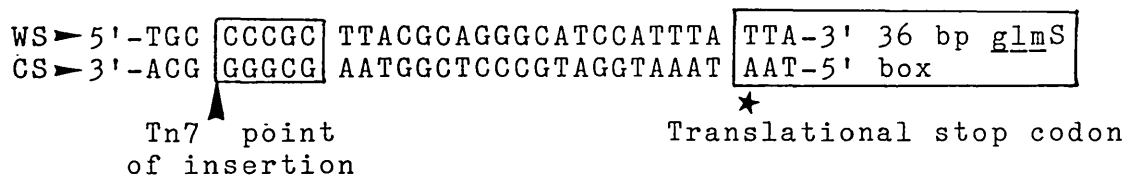
8.3 Use of Sodium Bisulphite as a Mutagen:

Strategies discussed above failed to yield single or multiple mutants with high efficiencies and this led to the use of sodium bisulphite as the mutagenic agent. Sodium bisulphite specifically deaminates deoxy cytidine to give deoxy uridine and this results in a CG to TA base-pair transition in an E. coli mutant deficient in uracil-N-glycosylase (ung⁻). Alternatively, the gap can be repaired in vitro using DNA polymerase, Klenow-fragment and the DNA transformed into a wild-type bacterial strain. Although, this method is restricted to the production of transition mutations of the type GC to AT, it has been used successfully (Berg et al., 1984; Shortle and Nathans, 1978; Smith, 1982).

The targets for this mutagenesis were 70 bp attTn7 DNA cloned in the M13 phage vector, mp10 (MAS) and the 89 bp attTn7 DNA (Δ L849) containing an additional 19 bp of the glmS sequence which were not required for wild-type transduction frequencies. The Δ L849 was obtained after the deletion analysis of the primary attTn7 DNA from the left end as discussed in chapter 4. The target DNA in this clone was also present in M13 phage vector, mp10, but in the reverse orientation to MAS. Since the 70 bp attTn7 fragment was cloned only in one possible orientation, Δ L849 was employed to generate the base substitution in the second strand of the attTn7. In the (+) strand of these two clones mutagenesis by sodium bisulphite can give rise

to a C to T mutation [corresponding to a G to A mutation in the (-) strand]. In order to avoid confusion, the (+) strands of MAS and Δ L849 (which are complementary) were designated as the "Watson" and "Crick" strands respectively in this study (see figure 8.3).

Since the sodium bisulphite treatment requires single-stranded DNA, GC to AT mutations in the attTn7 DNA were generated by treating a heteroduplex formed between a vector containing the attTn7 DNA and a vector DNA with a deletion in the attTn7 DNA sequence.



WS (+) strand of DNA in the M13 clone MAS, (referred to as the Watson strand of attTn7 DNA)

CS (+) strand of DNA in the M13 clone Δ L849, (referred to as the Crick strand of attTn7 DNA)

Figure 8.3 Watson and Crick strands of target attTn7 DNA for mutagenesis studies.

In these heteroduplexes the target attTn7 DNA remained the only single-stranded region and was then susceptible to the single-stranded mutagen, sodium bisulphite (see chapter 2, section 2.2.3B). This method generated point and multiple mutants with high frequencies (greater than 40%).

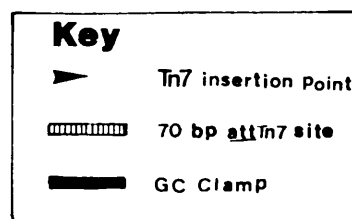
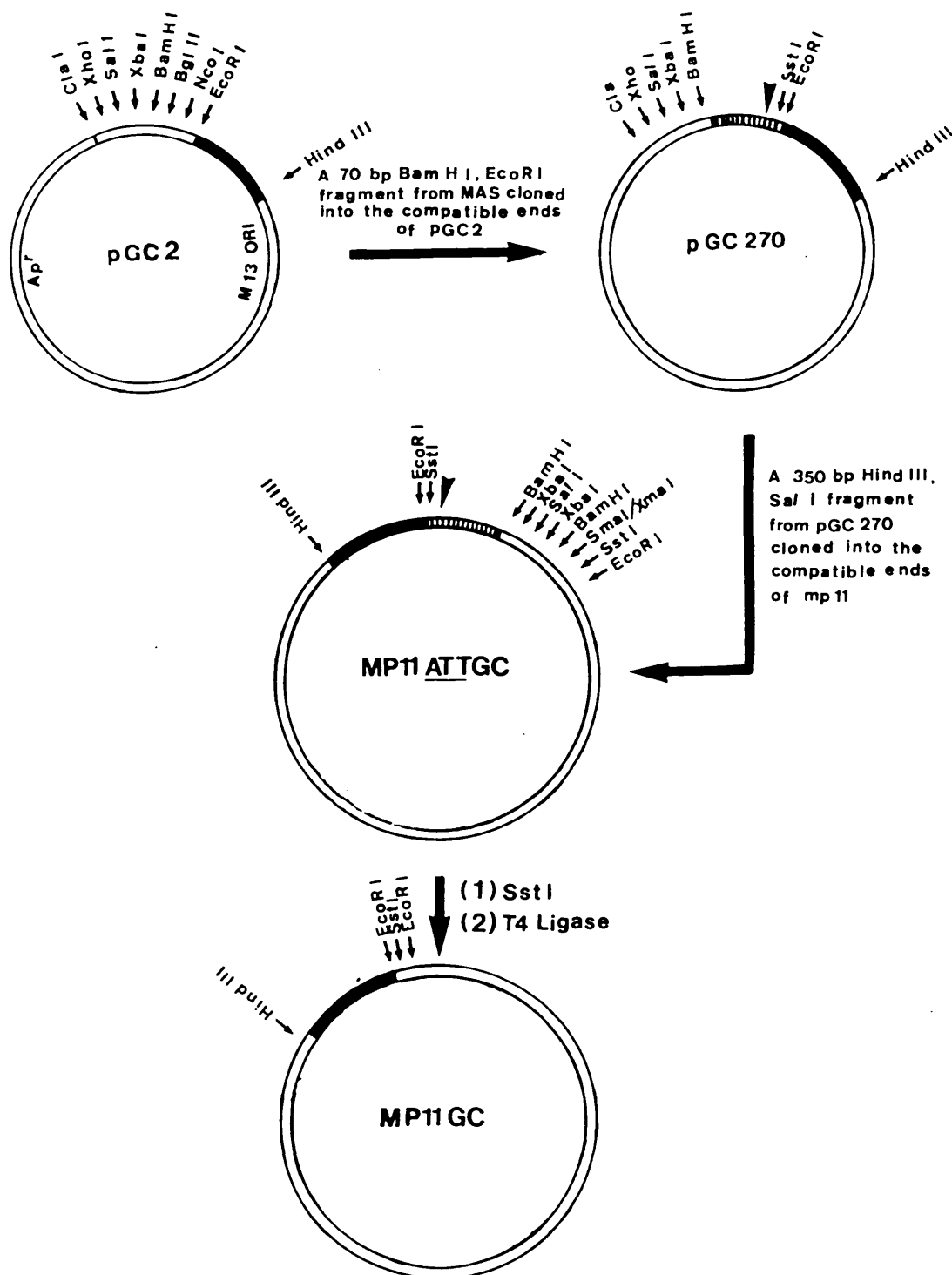


Figure 8.4: Strategy for construction of M13 phage vectors, MP11ATTGC and MP11GC for sodium bisulphite mutagenesis.

8.3.1 Alteration of C to T in the Watson strand of 70 bp attTn7 DNA (MAS):

This was achieved by mutagenesis in the M13 phage vector containing the 70 bp attTn7 DNA sequence and a 250 bp GC rich DNA (GC clamp) of plasmid pGC2 in a clone designated MP11ATTGC.

The strategy for construction of this clone is described in figure 8.4. This clone was designed to obtain mutants in the attTn7 DNA with sodium bisulphite treatment followed by running the Myers's gradient gels. Although the Myers's technique failed to recover mutants with high efficiency in the previous set of experiments it was hoped to optimise the yield of mutants using different chemical mutagenesis. For this reason, the M13 phage vector, MP11ATTGC was subjected to sodium bisulphite treatment rather than MAS. In order to obtain single-stranded attTn7 DNA as a target in the heteroduplex, it was essential to remove the 70 bp of the attTn7 DNA sequence from this clone. The M13 clone with a 70 bp deletion of attTn7 DNA in MP11ATTGC was constructed as described in figure 8.4 and named MP11GC.

0.2µg of MP11ATTGC DNA (single-stranded) and MP11GC (RF), EcoRI digested were mixed in 10mM MgCl₂ and 5mM NaCl in a 20µl volume. The mixture was boiled for 5 minutes at 95 °C and on cooling annealed to the opposite strand to obtain the desired heteroduplex molecule.

This procedure can produce three forms of DNA molecules, (i) heteroduplex formed between MP11ATTGC and MP11GC, with the target DNA as single-stranded region, (ii) single-stranded MP11ATTGC and (iii) reannealed linear RF form of MP11GC. This was confirmed by running a 0.8% agarose electrophoresis gel (data not shown). The form I was the target of mutagenesis which contained only the 70 bp attTn7 as a single-stranded DNA.

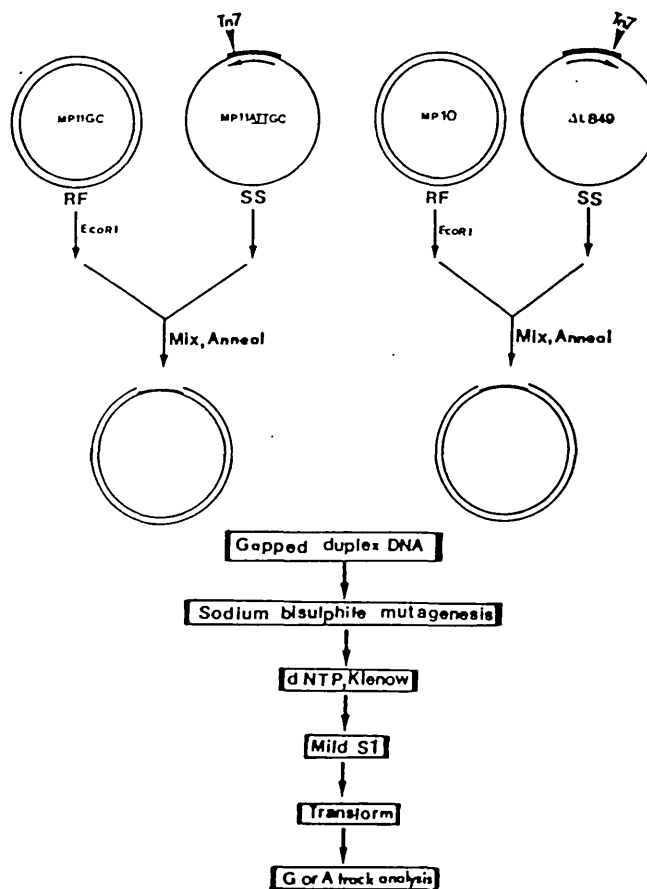


Figure 8.5 Strategy adopted for sodium bisulphite mutagenesis. The broader line shows the target DNA and arrows indicate the direction of inserts. The target for sodium bisulphite treatment in MP11ATTGC is 70 bp of attTn7 site and referred to as the Watson strand while in Δ L849 is 89 bp of attTn7 site and referred to as the Crick strand.

The background of single-stranded MP11ATTGC DNA and linear MP11GC (RF) were removed in two steps: first, by treating the DNA with S1 nuclease (after completion of the mutagenesis steps) and second, in the absence of T4 DNA ligase linear DNA upon transformation produces plaques with very low frequency. Therefore, it was not required to purify the heteroduplex DNA for mutagenesis. The procedure for sodium bisulphite mutagenesis is described in chapter 2, section 2.2.3B and the strategy adopted is diagrammatically shown in figure 8.5.

Since the saturation mutagenesis was carried out in M13 phage vectors, it was easy to identify mutant clones by DNA sequencing (ddGTP reaction only) (Sanger et al., 1977). The mutant clones had lost single or multiple bands in the region of attTn7 DNA. Figure 8.6 (A) shows the autoradiograph of the DNA sequencing reactions (ddGTP track only) of 24 mutants identified. One mutant clone (number 25) was picked at random in order to determine the complete nucleotide sequence and figure 8.6 (B) shows the autoradiograph of this mutant clone. No other DNA rearrangements in attTn7 were found in mutant number 25 to be associated with the use of this method, therefore, it was assumed that all other mutants have only G to A transitions. Because C to T transition mutants in the 70 bp attTn7 were obtained with high frequency, it was not necessary to identify the mutants by running a Myers's gradient gel. All mutants identified are shown in figure

8.8. However, mutant number 9 (C to G) in this list was obtained in experiments (to produce the ClaI RE site) described in chapter 7 section 7.2. and is included only for a comparison. The transduction frequency results for each mutant clone are discussed in section 8.5.

8.3.2 Alteration of C to T in the Crick Strand of 89 bp attTn7 DNA (Δ L849):

To obtain this class of mutants, Δ L849 single-stranded DNA was mixed and annealed to the M13 phage vector, mp10 (RF, EcoR1 RE digested) in order to obtain gapped circle heteroduplex molecules for sodium bisulphite mutagenesis as described in section 2.2.3B(iv). After chemical treatment the mutant clones were identified by the same procedure as described for MAS. In total 56 mutants were identified and figure 8.7 (i and ii) shows the autoradiograph of the DNA sequencing reaction (ddGTP reaction only) of these mutants.

Although the C to T transitions were made in the 89 bp the attTn7 (Crick strand), figure 8.9 shows the identified mutants only in the 70 bp attTn7 in 5' to 3' direction for comparison with the mutants of MAS.

The DNA sequence shown is therefore of the Watson strand and C to T changes in the Crick strand correspond to G to A changes in the Watson strand. After further comparison of these mutants, 19 were found to be in duplicate in the

Figure 8.6 Autoradiograph of DNA sequencing gels showing mutants of the Watson strand of attTn7.

A: ddGTP track run of 24 mutants. The substitution from guanosine to adenosine at specific positions resulted in disappearance of the bands at these locations in the synthesised strand. These substitutions correspond to the cytidine to thymidine in the plus strand of M13 phage vector, MP11ATTGC. The wild-type (w.t.) is shown for comparative purposes.

B: DNA sequencing gel of mutant number 25 showing two guanosine residues substituted with adenosine residues as specified by arrows.

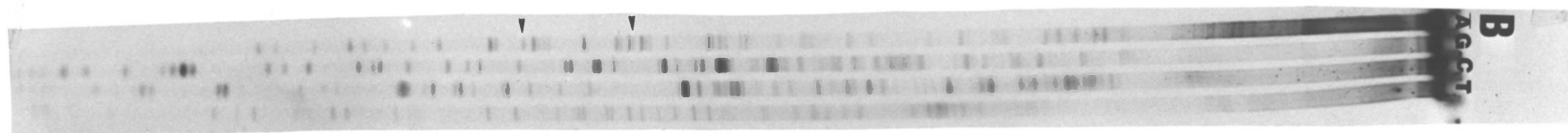
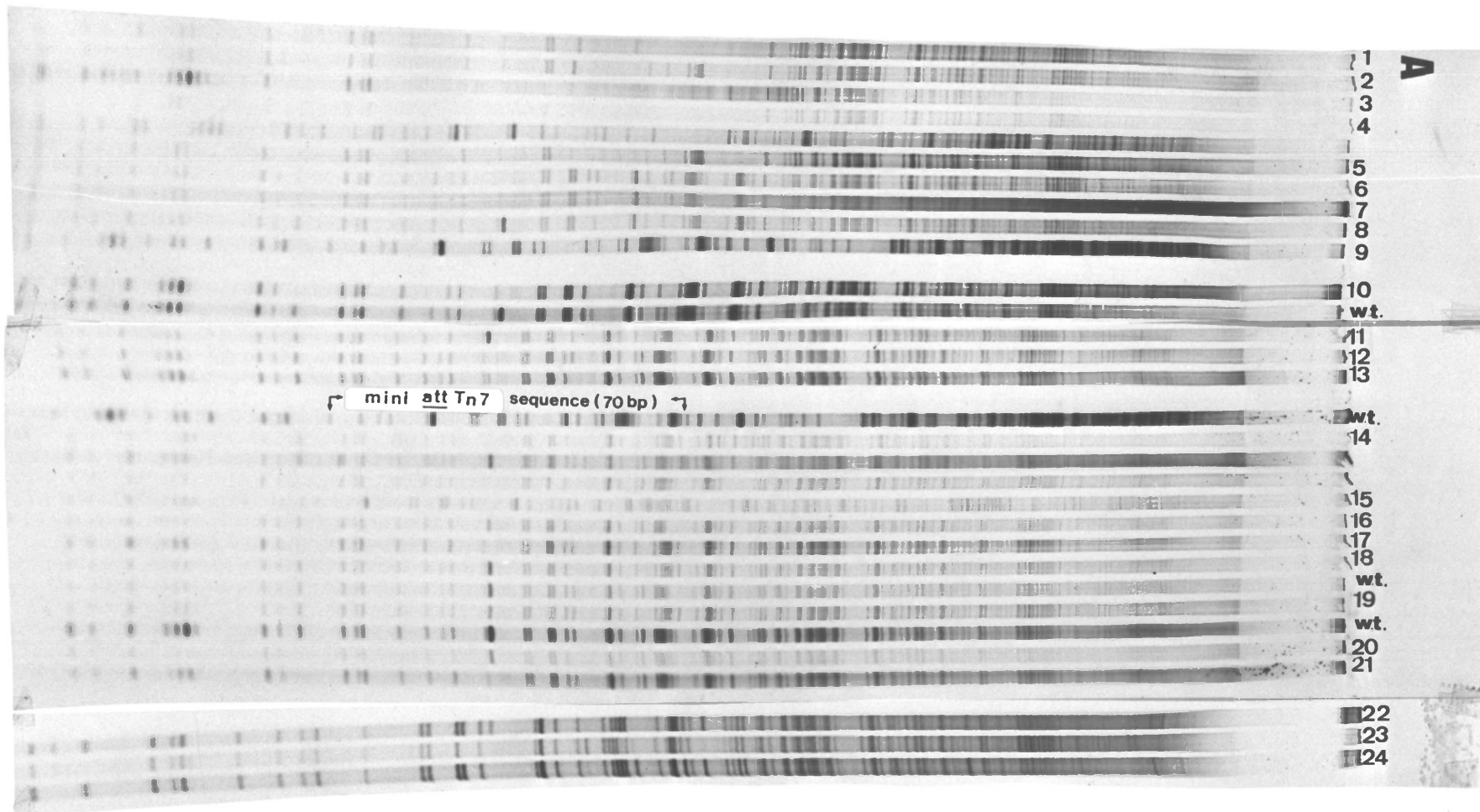


Figure 8.7(i) Autoradiograph of DNA sequencing gels showing mutants of the Crick strand of attTn7. The ddGTP track run of 29 mutants. The substitution from guanosine to adenosine at specific positions resulted in disappearance of the bands at these locations in the synthesised strand. These substitutions correspond to the cytidine to thymidine in the plus strand of M13 phage vector, Δ L849 The wild-type (w.t.) is shown for comparative purposes.

mini attTn7 site

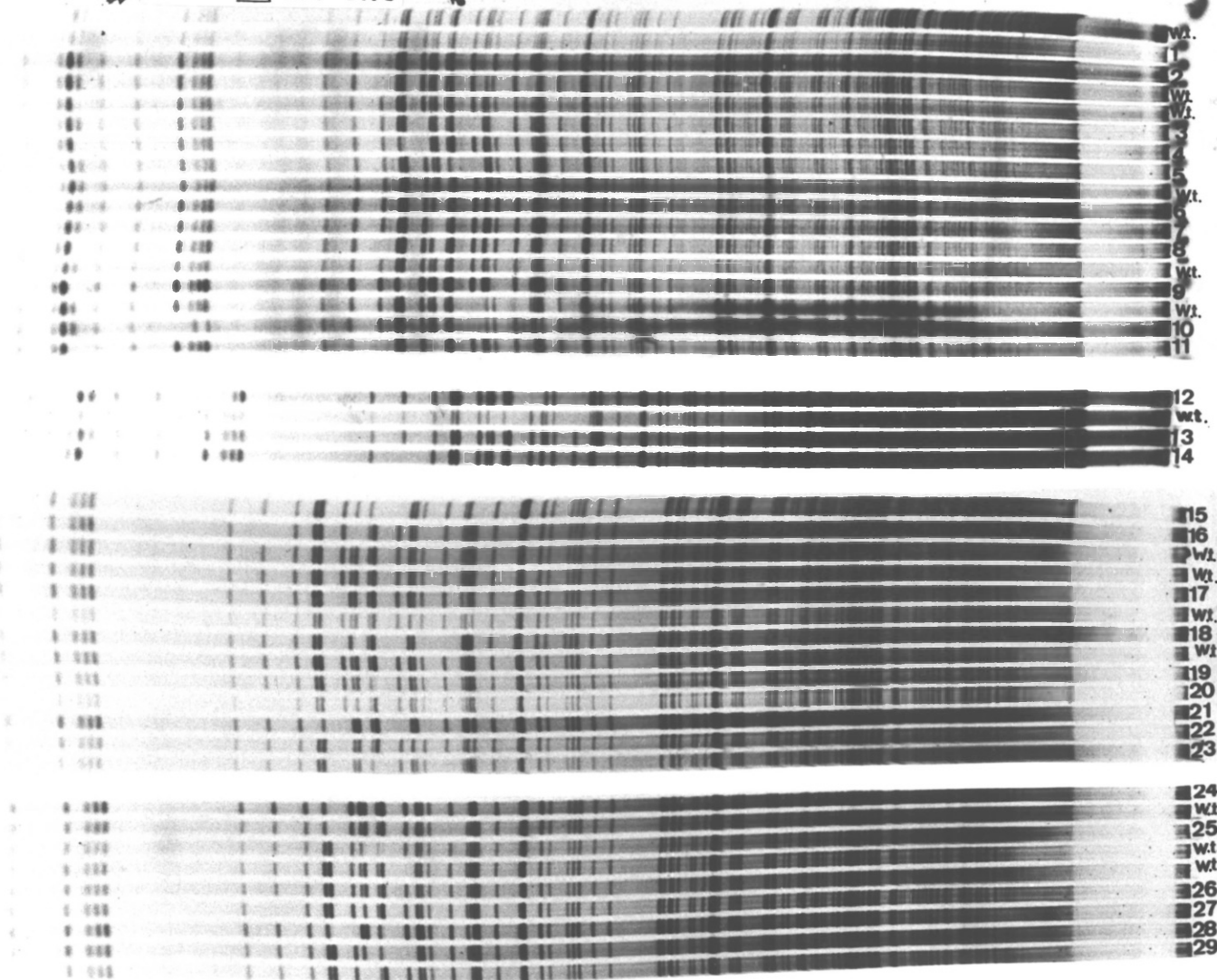
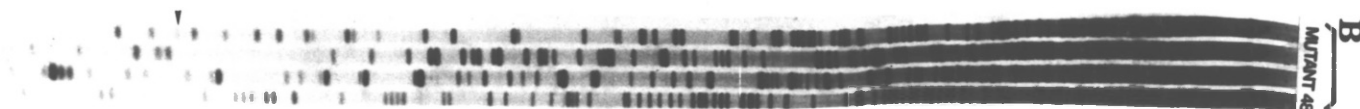
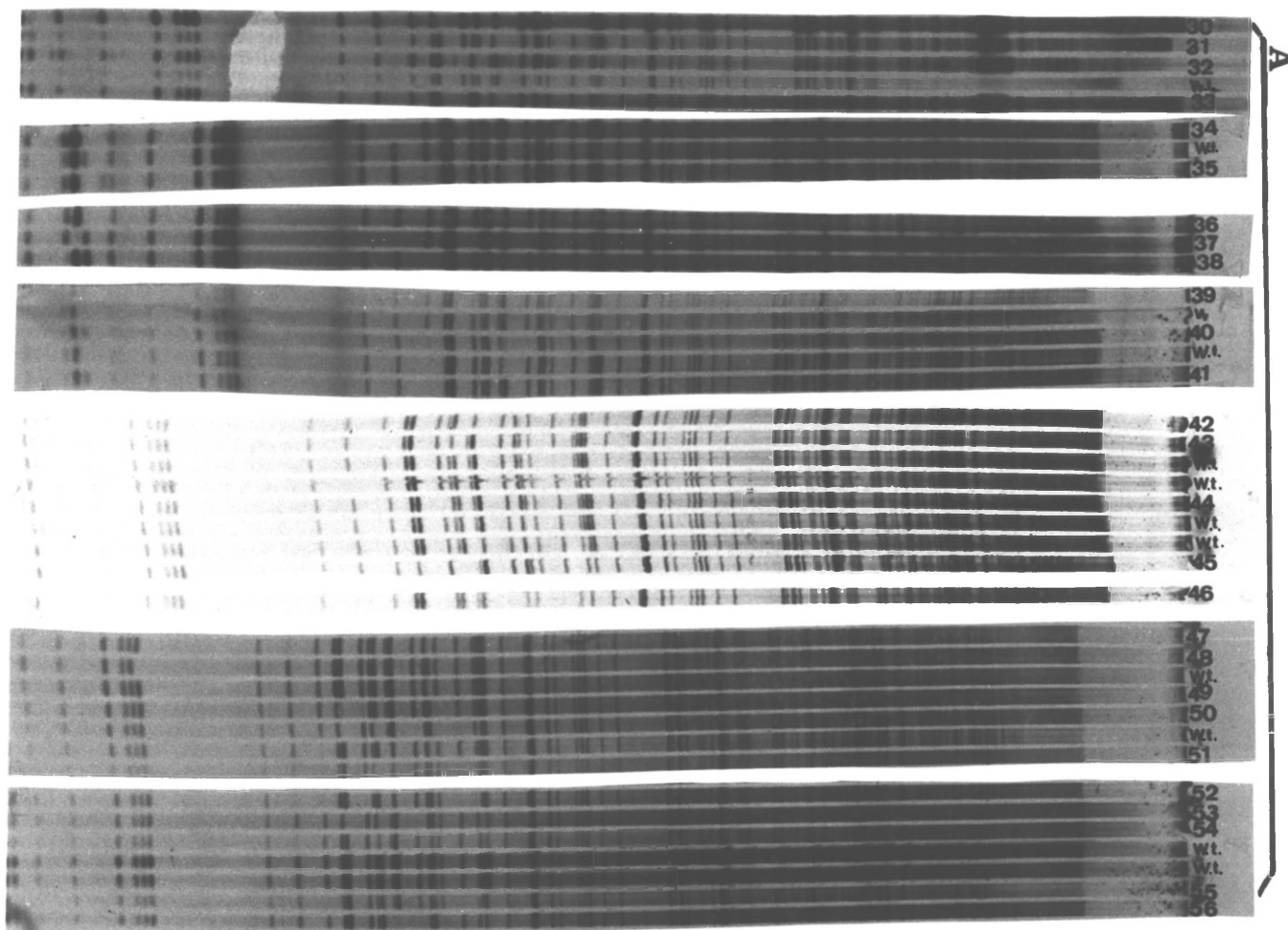


Figure 8.7(ii) Autoradiograph of DNA sequencing gels showing mutants of the Crick strand of attTn7.

A: The ddGTP track run of mutants (30 to 56). The substitution from guanosine to adenosine at specific positions resulted in disappearance of the bands at these locations in the synthesised strand. These substitutions correspond to the cytidine to thymidine in the plus strand of M13 phage vector, Δ L849. The wild-type (w.t.) is shown for comparative purposes.

B: DNA sequencing gel of mutant number 48 showing one guanosine residue substituted with adenosine residue as specified by an arrow.



region of 70 bp of attTn7 DNA, however these mutants have extra different base substitutions in the insert but outside the 70 bp region of interest. These mutants were still of some interest, therefore they were saved for further analysis by the M13 transduction frequency assay.

It was expected that base substitutions in the extra 19 bp in the Δ L849 would not cause any alteration in the transduction frequency. They are designated by an asterisk in figure 8.9. The transduction frequency results for each mutant clone are discussed in section 8.5.

8.4 Distribution of Base Pair Substitutions (C to T and G to A) in the 70 bp attTn7 DNA:

On compiling the mutants obtained from C to T or G to A mutation, a higher frequency of mutations was clearly observed in the 36 bp glmS box in the 70 bp attTn7 DNA.

The lower frequency of mutants found in the non-coding region of glmS may be due to the formation of secondary structure; a double-stranded DNA region constituting the stem. It is evident from the results particularly in the class of C to T mutants that 14 out of 24 base substitutions occur in the centre of the inverted repeat i.e., the unpaired loop of the cruciform which is a single-stranded DNA region. Since single-stranded DNA is more susceptible to sodium bisulphite treatment than double-stranded DNA, this distribution is not unexpected. The

mutants of both (C to T) and (G to A) classes can be further segregated according to the individual changes made in the 70 bp attTn7 DNA and these classes are shown in table 8.1.

Table 8.1 Distribution of mutants in the 70 bp attTn7 site.

Number of mutations	Number of mutants	Mutant number given in figures 8.8 and 8.9
C to T Class		
Single	2	7, 13
Double	6	6, 18, 19, 21, 22, 25
Triple	6	8, 14, 15, 16, 17, 24
Quadruple	3	10, 20, 23
>Quadruple	5	1, 2, 3, 4, 5
G to A Class		
Single	11	2, 3, 5, 6, 9, 10, 14, 34, 36, 47, 48,
Double	15	1, 4, 7, 8, 12, 13, 15, 18, 20, 21, 23, 24, 38, 43, 56
Triple	9	11, 16, 25, 27, 30, 35, 37, 45, 51
Quadruple	1	52

8.5 Discussion on the Results of the Mutagenesis Studies:

All mutants of both classes identified (i.e. C to T and G to A) were analysed as targets for transposition of Tn7 using the M13 transduction frequency assay. Tables 8.2 and 8.3 show the transduction frequency results for each mutant of C to T and G to A classes respectively. Each mutant clone was tested 2-3 times and the transduction frequency values were compared with the wild-type; these relative values are also presented in the tables 8.2 and 8.3.

<i>K. pneumoniae</i> (81 bp of attTn7)	5'-ATG CCTGC C- TGCCTCCGCTTCTCGGGACGCGGCGCG G..G..G.....C..G.....G.....-3'	Mean relative transduction frequency value for each mutant clone to the wild type (MAS), the results of individual experiments are shown in table 8.2.
<i>S. marcescens</i> (70 bp of attTn7)	5'-CCG GCTTT TGAAGTTTACTGCGACAAA T..G..G..A..C..C.....G.....-3'	
<i>E. coli</i> (70bp attTn7 (MAS))	5'-TGC CCGCG TTACGACGGGCATCCATTTA TTAACAACCGTAACCGAATTTGCCAGGTTACGCGGCTGGTC-3' 3'-ACG GGGCG AATGCGTCCCGTAGGTAAT AATGAGTTGGCATTGGCTAAACGCGTCCAAATGCGCGGACGAG-5' ←-----→ * DNA encoding the C-terminal domain of <i>glmS</i> ←-----→ <i>glmS</i> transcriptional terminator	1.0E 0
Mutant number		
1-	5'-..I TTT.I ...I.I.....I..... ..I.I...I...Π.....I.....I.....I.....-3	2.6E-5
2-	..I I.....I..Π..... ..I.I..I.....Π.....I.....I.....I.....I	1.2E-5
3-	..I TTT.I ...I.I.....I..Π..... ..I.I..Π.....Π.....Π.....I..I.....I	2.2E-5
4-	..I TTT.I ...I.I.....I..Π..... ..I.I..Π.....Π.....I.....I..I.....I	8.7E-6
5-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	3.5E-5
6-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	6.5E-3
7-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.4E-1
8-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.2E-2
9-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.4E 0
10-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.1E-2
11-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	3.3E-1
12-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	9.3E-2
13-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	2.7E-1
14-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	4.6E-2
15-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	8.0E-2
16-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	9.0E-2
17-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	7.0E-2
18-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	2.9E-1
19-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	2.7E-1
20-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	3.6E-2
21-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	4.9E-2
22-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.3E-1
23-	..I I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.2E-2
24-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	1.0E-2
25-	I.....I.....Π..... ..I.I..I.....I.....I.....I.....I.....I.....	3.7E-2

Figure 8.8 Results of mutagenesis studies on the Watson strand of attTn7 DNA. The (+) strand of MP11ATTGC was referred to as the Watson strand. In all cases the mutant clones were confirmed by DNA sequencing (ddGTP track analysis only) of the MAS DNA and the results are shown in the 5' to 3' direction. The mutant numbers 7 and 8 are identical to numbers 12 and 11 respectively, however they were identified independently. Mutant number 9 was generated by the method described in chapter 7, section 7.2 and this mutant is placed in the list only for comparison. The complete nucleotide sequence of mutant number 25 was determined and DNA sequencing autoradiograph is shown in figure 8.6 (B). The "T" in each lane represents the changes in the nucleotide C of the Watson strand of the 70 bp attTn7 DNA and the dots are the unchanged nucleotides. The DNA sequences of attTn7 sites of *E. coli*, *K. pneumoniae* and *S. marcescens* (1st site) are shown for a comparison. In *E. coli* attTn7 the 5 bp Tn7 target duplication is displayed in box and a star shows the translational stop codon of *glmS*.

Figure 8.9 Results of mutagenesis studies on the Crick strand of attTn7 DNA. The (+) strand of Δ L849 was referred to as the Crick strand of the attTn7 DNA in this study. Due to the cloning strategy used, the target site shown above had an additional 19 bp region whose nucleotides are not required for the wild-type transduction frequency (see chapter 4, section 4.3). There are 19 mutants in this figure which are identical in the 70 bp attTn7; however base substitution differs in the non-essential 19 bp region (data not included). These mutants are marked by an asterisk. Only 70 bp sequence is shown in 5' to 3' direction for simplicity. In all cases A represents the change in the corresponding G in the Watson strand of the 70 bp attTn7 DNA and the dots are unchanged nucleotides. The DNA sequence of attTn7 sites of E.coli, K.pneumoniae and S.marcescens (1st site) are shown for a comparison. In E. coli attTn7 the 5 bp Tn7 target duplication is displayed in box and a star shows the translational stop codon of glmS.

K. pneumoniae 5'-ATG [CTGC] C-
(81 bp of attTn7)
TCCCCGCTTCTCGGGACGCGGCGGCG G..G..G..C..G.....G.....-3'
(70 bp of attTn7)
S. marcescens 5'-CG [GCTT] TGAAGTTTACCTGCGACAAA T..G..G..A..C..C.....G.....-3'

Mean relative transduction frequency value for each mutant clone to the wild-type (MAS), the results of individual experiments are shown in table B.3

E. coli 5'-TGC [CCGCG] TTACGAGGGCATCCATTTA TTAAGTCAACCGTAACCGATTTTCCAGGTTACGCGCTGGTC-3'
[70bp
attTn7
(MAS)] 3'-ACG [GGGCG] AATGCGTCCGCTAGGTAAT AATGAGTTGGCATTGGCTAAACGCTCCATGCGCGGACCAAG-5'
← → * DNA encoding the C-terminal domain of *glsS*
glsS transcriptional terminator

1.0E 0

Mutant
Number

1.	5'-...A.....A.....-3'	2.1E-1
2.	A.....	2.9E-1
3.	A.....	1.1E-1
4.	A..A.....	1.0E-1
5.	A..	8.6E-1
6.	A.....	2.2E-1
7.	AA.....	1.7E-1
8.	A.....A.....	2.7E-2
9.	A.....	4.2E-1
10.	A.....	1.9E-1
11.	A.....AA.....	2.0E-2
12.	A.....A.....	8.5E-1
13.	A..A..	9.6E-2
14.	A.....	1.7E 0
15.	A..A..	1.1E-1
16.	A.....AA.....	5.0E-2
17.	A.....*	8.0E-2
18.	A..A..	5.6E-2
19.	A..*	1.2E 0
20.	A.....A.....	8.0E-2
21.	A.....A.....	2.3E-2
22.	A.....*	8.1E-2
23.	A.....A..	8.7E-2
24.	A.....A..	3.0E-1
25.	A.....A..A..	1.1E-1
26.	A.....A.....*	6.0E-2
27.	A.....A..A..	1.1E-1
28.	A.....A..A..*	1.6E-1
29.	AA.....*	1.6E-2
30.	A.....A..A..	3.2E-1
31.	A.....*	1.5E 0
32.	A..A.....*	4.2E-2
33.	A.....A..*	2.0E-1
34.	A.....	2.0E-1
35.	A.....A..A..	3.7E-1
36.	A.....	1.2E 0
37.	A.....A.....A.....	1.2E-2
38.	A.....A.....	1.2E-1
39.	A.....*	1.0E-1
40.	A.....A..*	1.2E-1
41.	A.....*	1.1E-1
42.	A.....	1.2E 0
43.	A.....A.....	1.0E-2
44.	A.....*	1.3E-1
45.	A.....AA.....	1.6E-2
46.	A.....A.....*	5.0E-2
47.	A.....	1.0E 0
48.	A.....	4.9E-1
49.	A.....A.....*	1.9E-1
50.	A.....A.....*	7.6E-2
51.	A.....A..A..	1.9E-2
52.	A.....A.....A..	6.3E-3
53.	A.....*	1.4E-1
54.	A.....A.....*	4.3E-2
55.	A.....*	2.1E-1
56.	AA..	1.0E 0

The effects of single or multiple base substitutions in attTn7 on the transduction frequency of Tn7 are evaluated in this study. Since many of the mutants obtained have an altered phenotype, it is possible to draw some conclusions on the contribution of individual nucleotides to the function of attTn7 DNA. To understand the effects of C to T or G to A mutants in the attTn7 DNA, mutations in the glmS 36 bp box and the glmS transcriptional terminator are evaluated separately. The mutations at particular positions are compared with the reference point of deletion mutants from the right end of attTn7 site; for example the first nucleotide of 36 bp glmS box corresponds to deletion mutant R26 and so on. However, the effects of multiple mutants in both domains of attTn7 can not be evaluated easily. The mean transduction frequency values of these mutants are listed for comparison in figures 8.8 and 8.9.

All mutants in the 36 bp glmS box having 3 fold or more negative effect on the transposition frequency are listed in table 8.4. Nevertheless these experimental values show an approximate fluctuation of ± 2 fold. Those mutants which have no clear effect on the transposition frequency are listed in figure 8.10 and include only 7 mutants of the G to A class (numbers 9, 14, 19, 31, 36, 42 and 56). In these mutants base substitutions occur at positions 22, 23, 34 and 51.

Table 8.2 Results of the M13 transduction frequency assays of the Watson Strand (C to T class mutants) attTn7 site. Their relative values to the wild-type (MAS) and mean relative values are also shown in the table.

Target	Independent trials						Mean relative value
	Transduction frequency			Relative value to MAS			
	1	2	3	1	2	3	
MAS	1.0E- 4	2.0E- 5	1.0E- 5	1.0E 0	1.0E 0	1.0E 0	1.0E 0
1'	2.0E-10	5.0E-10	5.1E-10	2.0E-6	2.5E-5	5.1E-5	2.6E-5
2'	1.0E-10	2.2E-10	2.5E-10	1.0E-6	1.1E-5	2.5E-5	1.2E-5
3'	2.1E-09	5.1E-10	2.1E-10	2.1E-5	2.5E-5	2.1E-5	2.2E-5
4'	5.1E-10	2.1E-10	1.1E-10	5.1E-6	1.0E-5	1.1E-5	8.7E-6
5'	1.0E-10	1.1E- 9	5.1E-10	1.0E-6	5.5E-5	5.1E-5	3.5E-5
6	2.1E- 7	1.5E- 7	1.0E- 7	2.0E-3	7.5E-3	1.0E-2	6.5E-3
7	1.0E- 5	3.4E- 6	1.5E- 6	1.0E-1	1.7E-1	1.5E-1	1.4E-1
8	1.3E- 6	2.8E- 7	1.0E- 7	1.3E-2	1.4E-2	1.0E-2	1.2E-2
9	3.3E- 4	2.0E- 5	3.2E- 5	3.3E 0	1.0E 0	3.1E-1	1.4E 0
10	1.0E- 6	2.6E- 7	1.0E- 7	1.0E-2	1.3E-3	1.0E-2	1.1E-2
11	1.8E- 5	2.5E- 6	7.1E- 6	1.8E-1	1.2E-1	7.1E-1	3.3E-1
12	8.6E- 6	1.7E- 6	1.1E- 6	8.6E-2	8.5E-2	1.1E-1	9.3E-2
13	2.0E- 5	2.3E- 6	5.0E- 6	2.0E-1	1.1E-1	5.0E-1	2.7E-1
MAS	1.5E- 4	5.0E- 5	2.5E- 5	1.0E 0	1.0E 0	1.0E 0	1.0E 0
14	1.0E- 5	2.0E- 6	1.0E- 6	6.6E-2	4.0E-2	4.0E-2	4.6E-2
15	2.0E- 5	1.5E- 6	2.1E- 6	1.3E-1	3.0E-2	8.4E-2	8.0E-2
16	1.5E- 5	5.0E- 6	1.8E- 6	1.0E-1	1.0E-1	7.2E-2	9.0E-2
17	1.0E- 5	4.0E- 6	-	6.0E-2	8.0E-2	-	7.0E-2
MAS	1.0E- 4	2.0E- 4	1.0E- 5	1.0E 0	1.0E 0	1.0E 0	1.0E 0
18	5.0E- 5	3.5E- 6	2.0E- 6	5.0E-1	1.7E-1	2.0E-1	2.9E-1
19	1.5E- 5	3.0E- 6	5.1E- 6	1.5E-1	1.5E-1	5.1E-1	2.7E-1
MAS	1.5E- 4	5.0E- 5	2.5E- 5	1.0E 0	1.0E 0	1.0E 0	1.0E 0
20	1.0E- 5	1.5E- 6	3.0E- 7	6.7E-2	3.0E-2	1.2E-2	3.6E-2
21	1.0E- 5	2.0E- 6	1.0E- 6	6.7E-2	4.0E-2	4.0E-2	4.9E-2
22	5.0E- 5	2.1E- 6	5.5E- 7	3.3E-1	4.0E-2	2.2E-2	1.3E-1
23	4.0E- 7	1.5E- 6	1.5E- 7	2.6E-3	3.0E-2	6.0E-3	1.2E-2
24	2.0E- 6	5.0E- 7	2.0E- 7	1.3E-2	1.0E-2	8.0E-3	1.0E-2
25	1.0E- 5	2.0E- 6	3.0E- 7	6.0E-2	4.0E-2	1.2E-2	3.7E-2

The values written are less than or equal to those mentioned in the table.

Table 8.3 Results of the M13 transduction frequency assays of the Crick Strand (G to A class mutants) attTn7 site. Their relative values to the wild-type (MAS) and mean relative values are also shown in the table.

Target	Independent trials						Mean relative value
	Transduction frequency			Relative value to MAS			
	1	2	3	1	2	3	
MAS	1.0E-4	1.7E-5	7.5E-5	1.0E 0	1.0E 0	1.0E 0	1.0E 0
1	2.4E-5	3.0E-6	1.7E-5	2.4E-1	1.7E-1	2.2E-1	2.1E-1
2	3.0E-5	4.6E-6	2.3E-6	3.0E-1	2.7E-1	3.0E-1	2.9E-1
3	2.5E-5	1.0E-6	3.0E-6	2.5E-1	5.8E-2	4.0E-2	1.1E-1
4	3.1E-5	2.0E-6	1.0E-5	3.1E-1	1.1E-1	1.3E-1	1.8E-1
5	1.0E-4	1.0E-5	7.6E-5	1.0E 0	5.8E-1	1.0E 0	8.6E-1
6	1.0E-5	5.1E-6	2.1E-5	1.0E-1	3.0E-1	2.8E-1	2.2E-1
7	1.0E-5	5.0E-6	1.0E-5	1.0E-1	2.9E-1	1.3E-1	1.7E-1
8	4.5E-6	3.5E-7	1.4E-6	4.5E-2	2.0E-2	1.8E-2	2.7E-2
9	4.2E-5	2.5E-6	5.3E-5	4.2E-1	1.5E-1	7.0E-1	4.2E-1
10	1.4E-5	2.0E-6	2.4E-5	1.4E-1	1.1E-1	3.2E-1	1.9E-1
11	4.0E-6	2.0E-7	7.7E-7	4.0E-2	1.1E-2	1.0E-2	2.0E-2
12	6.8E-5	1.0E-5	1.0E-4	6.8E-1	5.8E-1	1.3E 0	8.5E-1
13	1.0E-5	2.0E-6	6.0E-6	1.0E-1	1.1E-1	8.0E-2	9.6E-2
MAS	2.0E-5	3.0E-5	5.0E-5	1.0E 0	1.0E 0	1.0E 0	1.0E 0
14	3.0E-5	6.0E-5	8.0E-5	1.5E 0	2.0E 0	1.6E 0	1.7E 0
15	1.2E-6	5.6E-6	4.5E-6	6.0E-2	1.8E-1	9.0E-2	1.1E-1
16	1.1E-6	1.3E-6	2.7E-6	5.5E-2	4.3E-2	5.4E-2	5.0E-2
17	1.5E-6	3.5E-6	4.3E-6	7.0E-2	1.1E-1	8.6E-2	8.8E-2
18	1.0E-6	2.1E-6	2.5E-6	5.0E-2	7.0E-2	5.0E-2	5.6E-2
19	1.5E-5	8.0E-5	2.0E-5	7.5E-1	2.6E 0	4.0E-1	1.2E 0
20	1.4E-6	2.7E-6	-	7.0E-2	9.0E-2	-	8.0E-2
21	6.0E-7	1.0E-6	3.0E-7	3.0E-2	3.0E-2	6.0E-3	2.3E-2
22	1.5E-6	2.1E-6	5.0E-6	7.0E-2	7.0E-2	1.0E-1	8.1E-2
23	2.2E-6	2.5E-6	3.5E-6	1.1E-1	8.3E-2	7.0E-1	8.7E-2
24	1.0E-5	3.0E-6	-	5.0E-1	1.0E-1	-	3.0E-1
25	4.0E-6	3.1E-7	-	2.0E-1	1.0E-2	-	1.1E-1
26	1.0E-6	2.1E-6	-	5.0E-2	7.0E-2	-	6.0E-2
27	2.5E-6	3.1E-6	-	1.2E-1	1.0E-1	-	1.1E-1
28	3.1E-6	5.1E-6	-	1.5E-1	1.7E-1	-	1.6E-1
29	3.4E-6	5.0E-6	-	1.7E-1	1.6E-1	-	1.6E-1
MAS	1.0E-4	2.0E-4	-	1.0E 0	1.0E 0	-	1.0E 0
30	1.0E-5	1.1E-4	-	1.0E-1	5.5E-1	-	3.2E-1
31	1.0E-4	4.3E-4	-	1.0E 0	2.1E 0	-	1.5E 0
32	3.5E-6	1.0E-5	-	3.5E-2	5.0E-2	-	4.2E-2
33	2.5E-5	3.0E-5	-	2.5E-1	1.5E-1	-	2.0E-1
34	1.5E-5	5.0E-5	-	1.5E-1	2.5E-1	-	2.0E-1
35	5.0E-5	4.8E-5	-	5.0E-1	2.5E-1	-	3.7E-1
36	1.5E-4	2.1E-4	-	1.5E 0	1.1E 0	-	1.2E 0
37	1.5E-6	2.1E-6	-	1.5E-2	1.0E-2	-	1.2E-2
38	2.7E-5	1.9E-5	-	2.7E-1	9.5E-2	-	1.2E-1
39	1.0E-5	2.0E-5	-	1.0E-1	1.0E-1	-	1.0E-1
40	2.0E-5	1.1E-5	-	2.0E-1	5.5E-2	-	1.2E-1
41	1.3E-5	2.1E-5	-	1.3E-1	1.0E-1	-	1.1E-1
MAS	1.5E-5	3.0E-4	3.5E-4	1.0E 0	1.0E 0	1.0E 0	1.0E 0
42	2.5E-5	2.1E-4	-	1.6E 0	7.0E-1	-	1.1E 0
43	1.7E-7	3.1E-6	-	1.0E-2	1.0E-2	-	1.0E-2
44	1.2E-6	5.8E-5	-	8.0E-2	1.9E-1	-	1.3E-1
45	1.5E-6	7.1E-5	-	1.0E-1	2.3E-1	-	1.6E-2
46	1.0E-6	1.5E-5	-	6.6E-2	5.0E-2	-	5.8E-2
47	1.0E-5	4.6E-4	-	6.6E-1	1.5E 0	-	1.0E 0
48	1.0E-5	1.3E-4	1.4E-4	6.6E-1	4.3E-1	4.0E-1	4.9E-1
49	3.4E-6	5.1E-5	-	2.2E-1	1.7E-1	-	1.9E-1
50	1.3E-6	2.0E-5	-	8.6E-2	6.6E-2	-	7.6E-2
51	5.0E-7	2.0E-6	-	3.3E-2	6.6E-3	-	1.9E-2
52	1.4E-7	1.0E-6	-	9.3E-3	3.3E-3	-	6.3E-3
53	1.8E-6	5.0E-5	-	1.2E-1	1.6E-1	-	1.4E-1
54	8.0E-7	1.0E-5	-	5.3E-2	3.3E-2	-	4.3E-2
55	5.0E-6	3.0E-5	-	3.3E-1	1.0E-1	-	2.1E-1
56	2.0E-5	2.5E-4	-	1.3E 0	8.3E-1	-	1.0E 0

Table 8.4 Mutants in the 36 bp *glmS* box with negative phenotype. The mean relative values for each mutant clone to wild-type (MAS) are shown in ascending order.

Mutants of C to T and G to A classes ⁰		Mean relative values to wild-type (MAS)
w.t. 5'-TTACTCAACCGTAACCGATTTTGGCCAGGTTACGCGGCTGGTC-3'		1.0E 0 (1.0)
Point mutants		
22-	A.....*	8.1E-2 (12.3)
17-	A.....*	8.8E-2 (11.3)
12-	T..... ⁰	9.3E-2 (10.7)
39-	A.....*	1.0E-1 (10.0)
3-	A.....*	1.1E-1 (9.0)
44-	A.....*	1.3E-1 (7.6)
53-	A.....*	1.4E-1 (7.1)
7-	T.....	1.4E-1 (7.1)
34-	A.....	2.0E-1 (5.0)
6-	A.....	2.2E-1 (4.5)
13-	...T.....	2.7E-1 (3.7)
2-	A.....	2.9E-1 (3.4)
w.t. 5'-TTACTCAACCGTAACCGATTTTGGCCAGGTTACGCGGCTGGTC-3'		1.0E 0 (1.0)
Double mutants		
6-	TT.....	6.5E-3 (153.8)
43-	A.....A.....	1.0E-2 (100.0)
21-	A.....A.....	2.3E-2 (43.4)
8-	A.....A.....	2.7E-2 (37.0)
25-	...T.....T.....	3.7E-2 (27.0)
32-	A.A.....*	4.2E-2 (23.8)
54-	A.....A.....*	4.3E-2 (23.2)
21-	T.....T.....	4.9E-2 (20.4)
46-	A.....A.....*	5.0E-2 (20.0)
18-	A.A.....	5.6E-2 (17.8)
26-	A.....A.....*	6.0E-2 (16.6)
50-	A.....A.....	7.6E-2 (13.1)
20-	A.....A.....	8.0E-2 (12.5)
23-	A.....A.....	8.7E-2 (11.4)
13-	A.....A.....	9.6E-2 (10.4)
15-	A.....A.....	1.1E-1 (9.0)
38-	A.....A.....	1.2E-1 (8.3)
40-	A.....A.....*	1.2E-1 (8.3)
29-	AA.....*	1.6E-1 (6.2)
28-	A.....A.....	1.6E-1 (6.2)
7-	AA.....	1.7E-1 (5.8)
4-	A.....A.....	1.8E-1 (5.5)
49-	A.....A.....*	1.9E-1 (5.2)
33-	A.....A.....*	2.0E-1 (5.0)
24-	A.....A.....	3.0E-1 (3.3)
w.t. 5'-TTACTCAACCGTAACCGATTTTGGCCAGGTTACGCGGCTGGTC-3'		1.0E 0 (1.0)
Triple or quadruple mutants		
52-	A.....A.....A.....A.....	6.3E-3 (158.7)
37-	A.....A.....A.....A.....	1.2E-2 (83.3)
51-	A.....A.....A.....A.....	1.9E-2 (52.6)
11-	A.....AA.....	2.0E-2 (50.0)
20-	...T.....T.....T.....T.....	3.6E-2 (27.7)
16-	A.....AA.....	5.0E-2 (20.0)
15-	...T...T.....T.....	8.0E-2 (12.5)
16-	...T.....T.....T.....	9.0E-2 (11.1)
25-	A.....A.....A.....	1.1E-1 (9.1)
27-	A.....A.....A.....	1.1E-1 (9.1)
45-	A.....AA.....	1.6E-1 (6.2)
30-	A.....A.....A.....	3.2E-1 (3.1)
35-	A.....A.....A.....	3.7E-1 (2.7)

⁰ Mutant numbers are same as in figures 8.8 and 8.9.

* Mutants in the region of interest (70 bp *attTn7*) are identical, but extra different base substitutions exist outside this region.

⁰ Independent mutant identical to number 7.

() Fold drop in the calculated transposition frequency.

(ΔR) Base substitution positions correspond according to the deletion mutants as described in chapter 4, section 4.3.

Table 8.5 Mutants in the 36 bp glmS box with no effect on transduction frequency. The mean relative values for each mutant clone to wild-type MAS are shown in ascending order.

Mutants of G to A class θ		Mean relative values to wild- type (MAS)
w.t. 5'-TTACTCAACCGTAACCGATTTGCCAGGTTACGCGGCTGGTC-3'		1.0E 0
	$\Delta R51$ $\uparrow$ $\Delta R34$ $\uparrow$ $\Delta R22$ $\Delta R23$ $\uparrow$	
Mutants		
56-	AA..	1.0E 0
47-	A.....	1.0E 0
42-	A...	1.1E 0
19-	A..	1.2E 0
31-	A..*	1.5E 0
36-	A...	1.2E 0
9-	A.....	4.2E-1
5-	A..	7.7E-1

θ Mutant numbers are same as in figures 8.8 and 8.9.

* Mutants in the region of interest (70 bp attTn7) are identical, but extra different base substitutions exist outside this region.

(ΔR)_n Base substitution, positions correspond according to the deletion mutants as described in chapter 4, section 4.3.

The extensive and laborious analysis of mutant attTn7 DNAs is not conclusive because of the lack of A or T mutations. However, the G or C mutations in attTn7 were sufficient to draw the following conclusions about the involvement of nucleotides within the 36 bp glmS box for determining the high frequency transposition of Tn7.

Deletion analysis of attTn7 in the E. coli chromosome and comparison with attTn7 sites from K. pneumoniae and S. marcescens, has provided clues for the involvement of the DNA sequence encoding the C-terminal domain of glmS in the high frequency transposition of Tn7. The mutagenesis results have provided evidence for the role of individual nucleotides within the attTn7 whose structures are crucial in determining the fully functional activity of attTn7.

The nucleotide substitutions before the deletion point $\Delta R25$ and in the unpaired loop of the cruciform do not influence the transposition frequency; for example C to T class mutant number 10 has the same transposition frequency value as that of numbers 24 and 25 (see figure 8.8). The point mutants of C to T and G to A class at positions 25, 29, 38, 39, 45, and 58 show a drop of 4 to 12 fold in the transduction frequency, thus showing their involvement in the target site selection.

The results of double mutants can be analysed independently or can be compared with each other to make some conclusions. For example mutant number 6 has 2 base pair substitutions at positions 37 and 38 and show a drop of 150 fold in transduction frequency. Another example can be seen by comparing the mutants number 28 and 43 which have substitutions, one at identical position i.e. 27 and the other at positions 22 and 38 respectively. The combined effect of mutations at positions 27 and 38 is 16 fold higher than those of positions 27 and 22, thus suggesting that this drop must be associated with the mutation at position 38. It is also worth noting that alteration of 6 or more cytidine residues in the 36 bp glmS box has reduced the attTn7 site activity that of non-attTn7 site.

By comparing the all other mutants, it is possible to separate the 9 cytidine residues and 3 guanosine residues in the 36 bp encoding the C-terminal domain of glmS whose nucleotide structures are required for the wild-type

transduction frequency. These residues are marked by asterisks in the 36 bp glmS box (see figure 8.10). However no mutant was identified at position number 35.

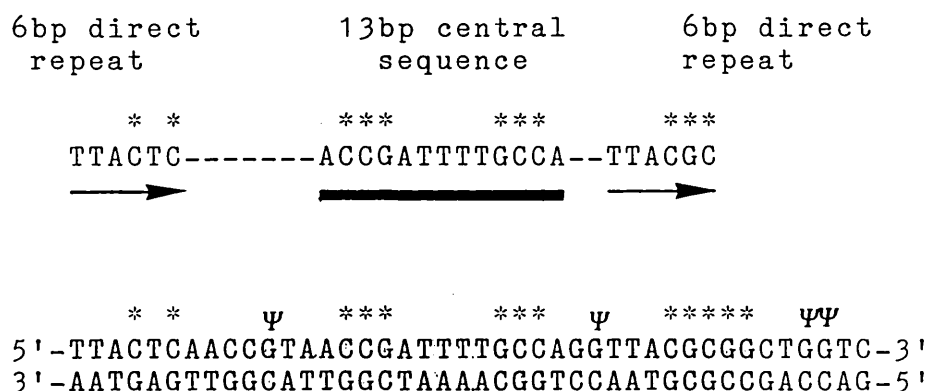


Figure 8.10 Important guanosine and cytidine residues in glmS 36 bp box. The glmS box is shown in bold typing and the sensitive nucleotides are marked by asterisks. The nucleotide substitutions with no effect on the transduction frequency of Tn7 are marked by ψ.

Three elements within the 36 glmS box of attTn7 can be identified by observing the effects of alteration of nucleotides C or G. The involvement of DNA sequences containing nucleotides A or T is only inferred from their linkage to DNA sequences containing C or G. Two 6 bp direct repeats with a consensus sequence of (5'-TAACNC-3') are separated by a total sequence of 22 bp which contains a 13 bp core sequence with a central dinucleotide TT and may be responsible for site-specific, high frequency transposition of Tn7. However to confirm the involvement of

these three possible elements within the 36 bp encoding the C-terminal domain of glmS, it would be necessary to examine the transition and transversion mutants in nucleotides A and T in the defined regions.

The experiments described in chapter 7 suggest the possible involvement of a secondary structure in attTn7 for high frequency target site selection. It is very difficult to infer any strong conclusion from the few mutants obtained in the glmS transcriptional terminator in the attTn7 site. Only 5 mutants were identified in the glmS transcriptional terminator of which three (numbers 10, 14 and 48) have point mutants of the G to A class. Another 2 mutants (number 23 of the C to T class and number 1 of the G to A class) have additional base substitution as shown in figures 8.8 and 8.9.

It can be concluded from this analysis that 13 nucleotides are involved in the attTn7 for high frequency transposition of Tn7. Any well defined role of the potential cruciform structure can not be documented on the basis of current information. Only 2 and 5 fold drop in the calculated transduction frequency in two point mutants number 10 and 48 (of C to T class) were observed which can be correlated with the thermostabilities of the glmS transcriptional terminator (see figure 8.11).

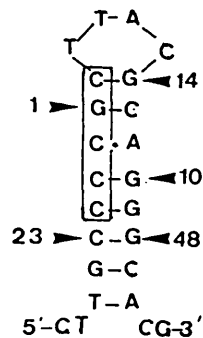
		ΔG Values (Kcal per mol)					
<u>glmS</u> transcriptional terminator		Wild- type	Mutant number 23	Mutant number 1	Mutant number 10	Mutant number 14	Mutant number 48
			C- T	G- A	G- A	G- A	G- A
		+5.0	+5.0	+5.0	+5.0	+4.0	+5.0
		-5.0	-5.0	-5.0	-5.0		-5.0
		-2.2	-2.2	+2.0	-2.2	-2.2	-2.2
		+2.0	+2.0	+2.0	+2.0	+2.0	+2.0
		-5.0	-5.0	-5.0	+2.0	-5.0	-5.0
		-5.0	-5.0	-5.0	-5.0	-5.0	-5.0
		-5.0	+2.0	-5.0	-5.0	-5.0	+2.0
		-2.2	-2.2	-2.2	-2.2	-2.2	-2.2
		-1.8	-1.8	-1.8	-1.8	-1.8	-1.8
		----	----	----	----	----	----
		-19.2	-12.2	-15.0	-12.2	-15.2	-12.2
Mean relative value of transduction frequency to that of wild- type (MAS)		→ 1.0E 0	1.2E-2	2.1E-1	1.9E-1	1.7E 0	4.9E-1

Figure 8.11 Mutants in glmS transcriptional terminator; their effect on ΔG values and correlation with transposition frequency. Thermodynamic (ΔG) values of the potential cruciform structure of the glmS transcriptional terminator in mutant attTn7 DNAs are calculated using the rule of Tinoco *et al.* (1973). Mutant numbers 1 and 23 have additional one and three base substitution respectively in the 36 bp glmS box in the 70 bp attTn7 DNA. The 5 bp target duplication is shown in a box. The mean relative transduction frequency value for each clone to the wild-type is shown. The asterisk shows that the mean relative value is linked with other mutations in mutants number 23 and 1.

8.6 In vitro DNA-Protein Interaction Studies:

These experiments were performed to study the interaction between the attTn7 DNA and Tn7 encoded proteins. The method used exploits the differential electrophoresis mobilities of free nucleic acid relative nucleic acid-protein complexes which are formed in vitro in an appropriate solution. The separation is achieved in a low ionic strength non-denaturing polyacrylamide gel. A typical electrophoresis pattern displays a diminution or disappearance of a DNA band and also, appearance of a low mobility band containing the specific DNA-protein complex (Strauss and Varshavsky, 1984; Wu and Crothers, 1984).

Whether a band of DNA-protein complex is seen after electrophoresis is dependent on many factors; the intrinsic stability of the complex, the salt concentration and the temperature within the gel, and the total time of electrophoresis.

8.6.1 Experimental Aspects:

The gel retardation assay was performed on the 70 bp attTn7 DNA which was obtained by cleaving the MAS (RF) DNA with HindIII and EcoRI. The sticky ends of the fragment were filled using radiolabelled dCTP and fractionated on a 0.8% non-denaturing gel so that the fragment could be isolated [see chapter 2, section 2.2.2B(vi)].

The Tn7 encoded proteins, identified by deletion analysis of Tn7 (Rogers et al., 1986), were not over-expressed in plasmid vector systems. Therefore, the total crude extract from two isogenic bacterial strains, one with Tn7 and one without, were used for this study. To obtain the highest level of Tn7 encoded proteins, a Tn7 donor replicon of a high copy number, pFK::Tn7 (pUC18 derivative) was chosen. The crude extract was prepared from both strains as described in chapter 2, section 2.2.4(i).

100ul of the extracts were loaded onto a 12.5% SDS-polyacrylamide gel as described in chapter 2, section 2.2.2.C(viii). No clear difference in the pattern of protein bands in the crude extracts of bacterial strains TG2 (pUC18) and TG2 (pFK::Tn7) was seen (data not shown).

The end filled DNA was mixed separately with varying amounts of crude extract from both strains, according to the conditions described in chapter 2 section 2.2.4(ii).

After incubation the mixture was loaded onto a 5% polyacrylamide gel and run at 20 mA, 4 °C, for 12 hours. The gel was autoradiographed and the results are shown in figure 8.12. The formation of a stable complex can be seen clearly by the slow mobility of attTn7 DNA in lanes 13, 14, 15 and 16, when incubated with the crude extract of strain TG2 (pFK::Tn7). This suggests that Tn7 encoded proteins have specific binding sites within the 70 bp of attTn7 DNA. This also shows that neither chromosomal nor pUC18 encoded

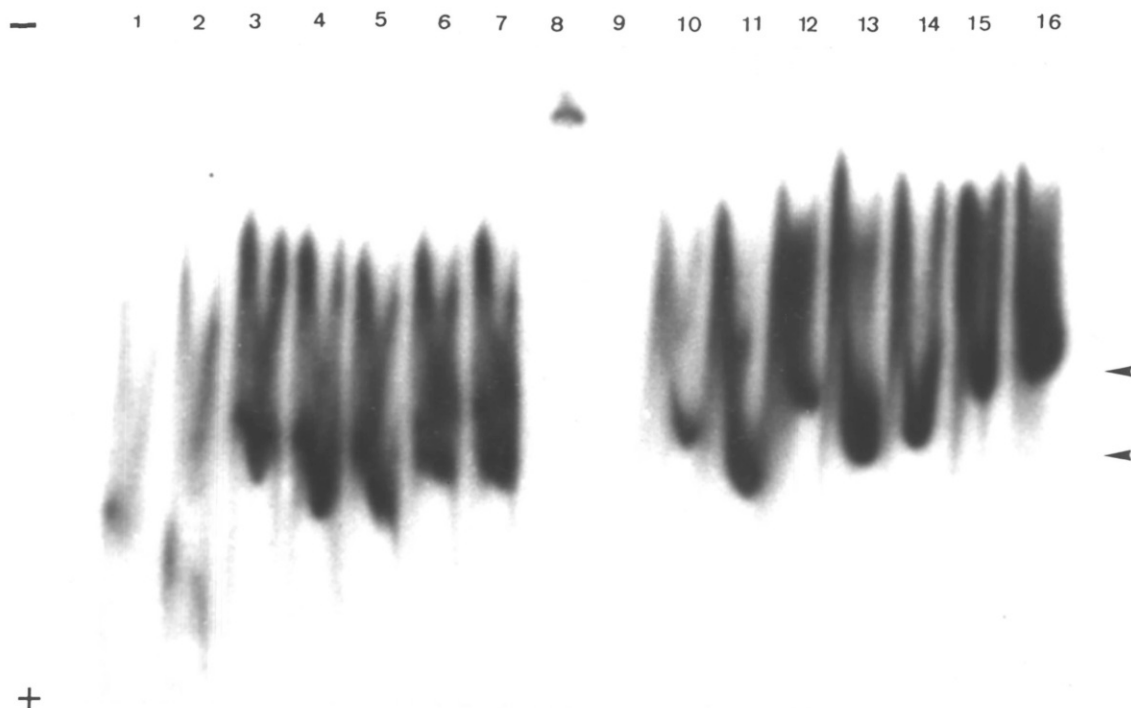


Figure 8.12 Binding gel showing 70 bp DNA fragment containing the attTn7. The gel used was 5% polyacrylamide and electrophoresed overnight in TE at 20 mA.

Lane 1: 10ng of HindIII and EcoRI 70 bp attTn7 fragment end filled with radiolabelled dCTP.

Lanes (2-7): 100ng of DNA as in lane 1 but incubated with 10, 20, 40, 60, 70 and 100 μ l of the crude extract of TG2 (pUC18) respectively.

Lane 8: linear MAS (7 Kb) DNA end filled with radiolabelled dCTP.

Lane 9: crude extract of the bacterial strain TG2 (pUC18).

Lanes (10-16): 100ug of DNA as in lane 1, but incubated with 10, 20, 40, 60, 70 and 100 μ l of crude extract of bacterial strain TG2 (pFK::Tn7) respectively.

It can be seen clearly that addition of crude extract of strain containing Tn7 in lanes 13, 14, 15 and 16 had an effect on the mobilities of the 70 bp attTn7 DNA.

proteins have affinity with the attTn7 DNA under the same conditions as seen in lanes 1-7.

The life time of the DNA-protein complex can be profoundly affected by the conditions of the electrophoresis. In figure 8.12, the DNA from the complex has dissociated in the run and has formed a smear of free DNA. It was difficult to obtain the optimal set of conditions using the crude extract and also due to the shortage of time, no further experiments were performed. However, theories and experimental evidence relating to the association and dissociation of DNA-protein complexes in vitro are available (Garner and Revzin, 1981; Strauss and Varshavsky, 1984; Wu and Crothers, 1984).

CHAPTER 9

Concluding Remarks and Future Experiments

9.1 Summary of Results:

The work presented in this thesis is the first study of the target sequence requirement for high frequency, site-specific transposition of a bacterial transposon. The insertional properties of other transposons were reviewed in chapter 1, section 1.4. Some transposons with random or less specific transpositional behaviour prefer AT rich DNA at the target site of the element [for example IS1 and Tn3 (Iida et al., 1982; Tu and Cohen, 1980)] or the mere presence of a GC base-pair at the ends of the target duplication sequence [for example Tn5 and Tn9 (Berg et al., 1984; Galas et al., 1980)]. Others have a reasonable consensus at the cleavage point which is duplicated to form a direct repeat at either end of the transposon [for example IS5 and Tn10 (Engler and Van Bree, 1981; Halling and Kleckner, 1982)].

Those transposons which display site-specific transposition such as IS4 (Klaer et al., 1980), ISH1 (Simsek et al., 1982), Tn554 (Murphy and Löfdahl, 1984) and Tn1545 (Cailland and Courvalin, 1987) have not been studied extensively; the DNA sequence analysis has been defined at

the specific insertion site but has not been analysed further.

The high frequency transposition of Tn7 into a unique chromosomal site in several bacterial species has provided the opportunity for the detailed analysis of the target site of Tn7 (attTn7) described in this thesis. A variety of approaches have been employed to probe the interaction between the transposon and its target site.

The foundation of this work was provided by the initial experiments of Lichtenstein (1980) and Lichtenstein and Brenner (1981; 1982), who reported the unique insertion of Tn7 in the chromosome of E. coli and were subsequently able to analyse the DNA sequence of the attTn7 site.

The experiments described in chapter 5 showed that Tn7 transposes with a high frequency into the chromosomes of Agrobacterium tumefaciens, Klebsiella pneumoniae and Serratia marcescens and into a unique site in the chromosome of A. tumefaciens and K. pneumoniae. These results extend the previous observations regarding the specificity of insertion of Tn7 in the chromosome of Escherichia coli (Lichtenstein and Brenner, 1981), Caulocacter crescentus (Ely, 1982), Pseudomonas aeruginosa (Caruso and Shapiro, 1982) and Xanthomonas campestris pv campestris (Turner et al., 1984).

The observation of two insertions of Tn7 in the chromosome of Serratia marcescens although unusual, is not unprecedented; there are reports of multiple insertion sites in the chromosome of a halotolerant, collagenolytic Vibrio strain (Thomson et al., 1981). In this bacterial strain most of the Tn7 insertions were shown to occur in the ilv operon all of which exhibited the same phenotype, suggesting, though not proving the existence of an attTn7 site. However, many lower frequency independent insertions were also reported to occur into the tryptophan, thymine, histidine and uracil operons located in the chromosome.

Deletion analysis of the primary E. coli attTn7 site DNA and the comparison with the attTn7 sites from the chromosomes of Klebsiella pneumoniae and Serratia marcescens (1st site) have revealed some striking features of Tn7 target site requirements.

Within the 64 bp E. coli attTn7 site, 36 bp encoding the C-terminal domain of glmS is required for high frequency transposition of Tn7. This demonstrates that the C-terminal coding region of the glmS gene has two functions. First, it provides a template for glmS mRNA transcription, and second, it is required for high frequency transposition of Tn7 into the adjacent transcriptional terminator. The nucleotide sequence in some other biological systems have multiple functions; for example the origin of phage lambda replication lies within the coding sequence of lambda protein O (Furth et al., 1977). However the requirement of

DNA sequence in the C-terminal region of a gene for the transposition of a bacterial transposon is unprecedented.

The DNA sequence comparison of the attTn7 sites from two additional bacterial chromosomes manifests the presence of a sequence which displays considerable homology with the DNA sequence of the right end of the E. coli attTn7 site. In all cases translation of the DNA sequence reveals a region of 100% homology in the predicted amino acid sequence to the C-terminal domain of glmS of E. coli. This suggests that in all cases high frequency transposition of Tn7 is directed into the 3' end of the glmS gene.

Glucosamine synthetase is an amido-transferase and is involved in an important cellular function; cell wall biosynthesis. The transfer of the (C-N) group from glutamine to fructose-6-phosphate leads to the synthesis of glucosamine. Other amido group transferases are involved in the synthesis of purines, pyrimidines, anthralate, p-aminobenzoic acid, histidine, asparagine, glutamate, glutamyl-t-RNA and nicotinamide adenine dinucleotide. The amino acid sequence in the region of the putative active site of glucosamine synthetase and other amidotransferases is shown in table 9.1 (Walker et al., 1984).

The comparison of different amido transferases suggests that the active site amino acids are located either at the N-terminal domain or the middle of the protein. In no case does the amino acid homology lie within the C-terminal

domain. The homology of attTn7 sites occurs in the C-terminal domain of glmS and is possibly involved in the function of the glmS gene product.

Table 9.1 Comparison of different amido-transferases at the active site amino acid residues. Taken from Walker et al. (1984).

Protein	Residues	Sequence
<u>E.coli</u> <u>glmS</u>	1-15	M C G I V G A I A Q R D V A E
<u>E.coli</u> <u>PRPP-AT</u>	1-14	C G I V G I A G V M P V N Q
<u>B.subtilis</u> <u>PRPP-AT</u>	1-14	C G V F G I W G H E E A P Q
<u>E.coli</u> <u>CPS</u>	261-276	T D I P V F G I C L G H Q L L A
<u>S.cervisiae</u> <u>CPS</u>	256-271	D C I P I F G I C L G H Q L L A
<u>E.coli</u> <u>ASII</u>	75-90	G K L P I I G I C L G H Q A I V
<u>S.marcescens</u> <u>ASII</u>	75-90	G R L P I I G I C L G H Q A I V
<u>E.coli</u> <u>PABASII</u>	71-86	G R L P I L G V C L G H Q A M A
<u>P.pudita</u> <u>aaASII</u>	71-86	G K L P I L G V C L G H Q S I G
<u>N.crassa</u> <u>ASII</u>	96-111	G K I P I F G V C M G Q Q C I F
<u>S.typhimurium</u>	-	A L G V C
Chicken <u>FGAR-AT</u>	-	G V C D B C Q

Where PARP-AT is glutamine phosphoribosylpyrophosphate amido transferase; CPS is a small sub-unit of carbamyl phosphate synthetase; ASII is anthranilate synthetase component II; PABASII is p-aminobenzoate synthetase; and FGAR-AT is formylglycinamide ribonucleotide amido transferase.

The transposition of Tn7 in the second attTn7 site of S. marcescens has occurred into a DNA sequence which does not display homology with the attTn7 sites of E. coli, K. pneumoniae or S. marcescens (1st site). The transposition of Tn7 in this sequence may involve a different mechanism. Whatever is the case; three key questions can be addressed on the basis of the results obtained in chapters 4, 5 and 6.

- (i) Why does Tn7 insert at a specific site in the chromosomes of many bacteria?
- (ii) Why is the insertion in the 3' end of the glmS gene?
- (iii) What governs the choice between the insertion in 3' region of the glmS gene and within the second attTn7 site in the S. marcescens genome?

The most obvious potential advantage of site-specific insertion is that the integration into a specific target site may be harmless to the host if that site is silent, whereas some events of random insertion will be harmful or lethal to the cell. Silent target sites could be either in transcribed DNA or in untranscribed DNA. In most cases the sequence of untranscribed DNA is not well conserved due to the lack of evolutionary selection, hence a given target sequence will not be maintained to provide a binding site for transposon encoded proteins. Transcribed DNA may provide a well conserved site if the sequence is important for the function of the gene or gene product but the gene would inevitably be inactivated by the insertion. For example the insertion element IS5 has been localised within the C-terminus (position 207 of the 232 codon) of the cI repressor gene of bacteriophage lambda (Kröger and Hobom, 1982).

Therefore, the ideal target for site-specific transposition of transposons and site-specific integration of bacteriophages or even plasmids would include a highly

conserved sequence that could direct insertion into an adjacent site where no critical function was affected. A conserved sequence at the amino- or carboxy-terminal of the coding region that allows insertion to the 3' or 5' end of the gene fits this criterion.

The transposition in the 5' end of a gene may occur, such as one insertion of IS2 in the ampC beta-lactamase gene of E. coli is reported to cause a 20% increase in the transcription level which is achieved by generating a new and highly efficient promoter in place of the previous one (Jaurin and Normark, 1983). A repetitive 340 bp sequence called Sigma has been found at the 5' end of many tRNA genes in Saccharomyces cerevisiae which apparently does not inactivate the gene function (Del Ray, 1982). Although the activation of genes by transposable elements has been observed in some cases (Galansdorff et al., 1980; Pilacinski et al., 1977; Reynolds et al., 1981), the insertion of transposable elements in the 5' end of a gene can also inactivate the gene function by disturbing the promoter sequence.

Site-specific integration of several bacteriophages may fit these same criteria. For example, satellite bacteriophage P4 and temperate phage lambda integrate at a unique site in the E. coli chromosome; at the 3' end of leucine tRNA synthetase gene and at the 5' end of biotin operon respectively. The integration of P4 (with the help of P2 proteins) in the chromosome is ensured in such a way that

it does not alter the expression of tRNA synthetase (Pierson and Kahn, 1987). However, there is no plausible explanation for the presence of a conserved domain (attB) in the non-transcribed region of E. coli required in site-specific recombination steps for integration of bacteriophages.

The answer to the second question may be that Tn7 has found the 3' end of the glmS gene to be an ideal target site on the criteria mentioned. It utilises a conserved DNA sequence at the C-terminal region of a gene whose amino acid sequence is 100% homologous in three bacterial species, thus suggesting its functional importance to the bacteria.

By recognising this highly conserved sequence, consequential insertion of Tn7 into the adjacent untranslated region is unlikely to be harmful to the organism. . It is possible that Tn7 insertion confers an advantage to the bacteria. It is observed that insertion of Tn7 into the chromosome of A. tumefaciens, K. pneumoniae and S. marcescens does not disturb any obvious metabolic function as the post transpositional colonies remain prototrophic. No visible difference was observed in the growth rate of bacterial colonies of the three species containing Tn7 or without Tn7 when grown on minimal glucose plates (this thesis).

The extent of the nucleotide sequence homology in the 36 bp encoding the C-terminal domain of the glmS gene among three bacterial species is described below:

<u>E. coli</u> / <u>S. marcescens</u>	30/36 (83%)
<u>E. coli</u> / <u>K. pneumoniae</u>	29/36 (80%)
<u>K. pneumoniae</u> / <u>S. marcescens</u>	33/36 (91%)

The DNA sequence in the 36 bp glmS box is highly conserved in comparison to the non-coding region where no DNA homology was found. Therefore it can be suggested that the Tn7 transposase recognition site lies within the coding region sequence. Tn7 encoded protein (probably transposase tnsD) and the 36 bp glmS box of different bacteria may have evolved independently, but for a very dependent relationship. However genetic mapping studies in Caulobacter crescentus suggest that Tn7 transposition occurs in a closely linked region of flaA-argA (Ely, 1982). The DNA sequence analysis of C. crescentus attTn7 would be of interest for a comparison with the known E. coli, K. pneumoniae and S. marcescens (1st site) attTn7 sites.

Gay et al. (1986) have proposed a model for Tn7 transposition regulation by glmS transcription proceeding into the right end of Tn7. They found a run over glmS transcript within the tnsA gene encoded by the right end of Tn7. This transcriptional read through may occlude transcription from a site within Tn7 since all five transposition genes are arranged in the same orientation

(Rogers et al., 1986). The transcription of tnsA may be linked with the transcription of other transposition genes. This may provide one control point for the expression of tnsA under different host cell conditions; growth rate dependence of glmS transcription which provides a link between cellular function and Tn7 transposition. Gay believes that this may inhibit the transposition of Tn7 from the chromosome to other replicons during rapid cell growth.

To test whether Tn7 insertion is beneficial to the cell or if it is just a genetic load on the cell, a study of the competitive experimental growth rates with and without Tn7 in a non-selective environment would be useful. If Tn7 insertion is advantageous for the cells, then the cells with Tn7 should out-compete the cells without Tn7. Beneficial growth rate effects in continuous culture have been reported for lysogens of bacteriophages lambda (Lin et al., 1977) and Mu, P1 and P2 (Edlin et al., 1977). Transposable element IS50R has also been shown to improve the growth rate of E. coli presumably by inhibiting the function of some growth inhibitor factor(s) (Hartl et al., 1983).

Tn7's genetic organisation is more complex than other transposons (Rogers et al. 1986; Carlos Flores, personal communication). Transposition of Tn7 into a preferred hot target site and into the non-preferred target site requires two different Tn7 encoded proteins tnsD and tnsE

respectively in association with three other proteins (tnsA, tnsB, tnsC) which are required in both modes of transposition. The DNA sequence analysis of the transpositional genes (Carlos Flores, personal communication) has revealed that some of the potential tns reading frames overlap slightly. This could indicate translational coupling between tns genes and is a situation most consistent with the hypothesis that tns genes are contained in two operons. The interaction of these five gene products for transposition of Tn7 in different bacterial host strains may involve a different relationship depending upon the genetic and physiological state of the cells.

The third question regarding the two insertions of Tn7 in the chromosome of S. marcescens cannot be answered easily. However, the second attTn7 site in the chromosome of S. marcescens may be due to the involvement of some cellular factor(s) with the tnsE dependent transposition or the second attTn7 site may represent a preferred site of insertion by the tnsE dependent pathway never before observed. It could also be explained that during cell growth the second event of transposition of Tn7 may have been triggered by the physiological state of the organism, as the bacterial colonies were treated repeatedly with the mutagen acridine orange. However, the induction of a second transposition event does not explain why the site has no similarity to the other attTn7 site.

9.2 Nature of the attTn7 DNA for Target Sequence Selection:

Particular stretches of DNA within attTn7 DNA which may be the recognition site of Tn7 encoded protein(s) (transposase) or host encoded proteins, were identified by random mutagenesis rather than by defining the contact points between the transposase and the target DNA by protection experiments.

With the aim of demonstrating a direct correlation between the attTn7 target site reactivity and the strength of protein binding using mutated and wild-type attTn7 DNA, foot print experiments were delayed until the mutational analysis was completed. The mutagenesis experiments (only G and C transitions) suggest the involvement of certain residues in the 70 bp attTn7 site for target site selection and high frequency transposition. These mutations led to altered phenotypes in the M13 transduction frequency assay, thus indicating their importance in transposition (possibly the recognition of the attTn7 site by Tn7 transposase). The base substitution of C to T or G to A in three defined regions within the 36 bp glmS box had a clear effect on the transposition frequency of Tn7 into the 70 bp attTn7 DNA. Two 6 bp direct repeats and a central 13 bp sequence in the 36 bp glmS box may be involved directly in the formation of a synaptic complex at the attTn7 site. The DNA sequence homology within these three domains of E. coli

Figure 9.1 DNA sequence comparison at three domains within the 36 bp of glmS box of E. coli (E.c) attTn7 with that of K. pneumoniae (K.p) and S. marcescens (S.m) (1st site) attTn7 sites. Positions marked with ϕ have mutations with altered effect on the transposition frequency. Where (*) indicates the mismatches within the three identified domains of the 36 bp of glmS C-terminal region which are thought to participate in the reactivity of the attTn7.

2nd direct repeat	13 bp central repeat	1st direct repeat
5'-TTACTC---(7bp)---	$\phi\phi\phi$ ACCGATTTTGCCA-(2bp)-	$\phi\phi\phi$ TTACGC-3' (<u>E.c</u>)
	* * *	*
5'-TTACTC---(7bp)---	ACCGACTTGCCA-(2bp)-	TTGCGC-3' (<u>K.p</u>)
	* * *	*
5'-TTACTC---(7bp)---	ACAGACTTCGCCA-(2bp)-	TTGCGC-3' (<u>S.m</u> 1st)

attTn7 site with the attTn7 sites of K. pneumoniae and S. marcescens (1st site) is outlined in figure 9.1.

The minimum DNA sequence involved from these three domains for wild-type target site selection (i.e. recognition of the Tn7 encoded transposase) can not be defined precisely unless the effects of base substitutions of A and T nucleotides in these regions are studied. Clearly the mutagenesis experiments alone are not sufficient to suggest the involvement of these domains in an intasome or synaptosome-like complex.

9.3 Synapsis at the attTn7 DNA and the Strand Cleavage at the Point of Insertion:

How do Tn7 encoded proteins bind to the 36 bp of the glmS gene (possibly within the three domains identified by mutagenesis studies) and proceed to cleave the DNA at sites of variable distance but an integral number of helical turns away? Many different models can be proposed to explain this mechanism.

All of these probably require that tnsD (transposase) interacts within the 36 bp glmS box in the attTn7. The transposition reaction steps that follow lead to the insertion of Tn7 DNA either 20 or 31 bp away from the 3' of the first nucleotide of the second repeat domain (5'-AATCGC-3') of attTn7.

On the basis of the results obtained from this work one conclusion to be drawn is that the attTn7 site consists of one specific domain and the point of insertion of Tn7 is located 3' of this sequence. It can also be inferred from this work that the DNA sequence at the point of insertion does not contribute significantly to target site selection. The mechanism of interaction between the different domains of the attTn7 can be explained to a certain extent, but clearly these models would be of a speculative nature. It is likely that the mode of action of Tn7 transposase is typical of Type III restriction enzymes.

Most site-specific enzyme-DNA interactions involve two separate events; first DNA sequence recognition, and second, the expression of catalytic activity. Different levels of complexity are associated with different enzymes on the basis of their substrate requirements and size (Boyer and Roulland-Dussoix, 1969). All known restriction endonucleases and methylases can be classified into three groups in accordance with the nomenclature proposed by Smith and Nathans, (1973); (i) Type I, require ATP, S-adenosyl-methionine plus magnesium, cleave or methylate the DNA usually 100-1000 bp away from the recognition sequence, (ii) Type II, require only magnesium, recognise a short nucleotide sequence on the DNA and cleave or methylate (separate enzymes) either within the recognition site or few base-pairs away from it, and (iii) Type III have intermediate properties between Type I and Type II, cleave or methylate a sequence at a fixed position (25-27 bp) away from the recognition site. There are only four known members of Type III restriction endonucleases and these include; (i) EcoP1, (ii) EcoP15 encoded by P1 bacteriophage and the related plasmid p15B, (iii) HineI and (iv) HinfIII from Haemophilus influenzae serotype e and f respectively.

Table 9.2 shows the restriction behaviour of some Type I, Type II and all Type III enzymes which recognise particular sequences and cleave a certain number of nucleotides away from this sequence. The Type I enzyme EcoK

Table 9.2 Restriction pattern of some of the Type I, II and III enzymes. The recognition sequence (RS) is shown in bold typing and the modification/cleavage site is shown 3' to the RS by (N)x, where x = number of nucleotides away from the RS. In Type I this number could be 100 to several thousand base pairs.

Enzyme	Recognition and cleavage sequence	Reference
Type I		
<u>EcoA</u>	5'-GAG-(N)7-GTCA-3'-(N)n	Bickle (1982)
<u>EcoB</u>	5'-CTCC-(N)7-CAGT-3'-(N)n	Bickle (1982)
<u>EcoK</u>	5'-AAC-(N)6-GTGC-3'-(N)n	Burchardt <u>et al.</u> (1981)
Type II		
<u>BbvI</u>	5'-GCAGC-3' --(N)8 --(N)12	Kessler and Holtkle (1986)
<u>FokI</u>	5'-GGATG-3' --(N)9 --(N)13	Kessler and Holtkle (1986)
<u>HgaI</u>	5'-GACGC-3' --(N)5 --(N)10	Brown and Smith (1977)
Type III		
<u>EcoPI</u>	5'-AGACC-3'-(N)25-27	Bachi <u>et al.</u> (1979)
<u>EcoP15</u>	5'-CAGCAG-3'-(N)25-27	Hadi <u>et al.</u> (1979)
<u>HinfIII</u>	5'-CGAAT-3'-(N)25-27	Bickle (1982)
<u>HineI</u>	5'-CGAAT-3'-(N)25-27	Piekarowicz <u>et al.</u> (1981)

is a large protein of 450 Kdal and consists of three sub-units; (i) R sub-unit (135 Kdal) required for restriction, (ii) M sub-unit (62 Kdal) required for modification, and (iii) S sub-unit required for sequence recognition. This enzyme has the capacity to cleave or methylate the DNA which could be several thousand base-pairs away from the recognition site and probably involves DNA translocation (Burchardt et al., 1981).

In Type III enzymes, for example EcoP1 and EcoP15, the distance between the recognition and the cleavage site is fixed (i.e. always 25 to 27 bp) and is of the order of 100Å. The protein is composed of two sub-units (i) modification sub-unit (106 Kdal) and the recognition sub-unit (73-76 Kdal) (Hadi et al., 1983). Therefore the holoenzyme seems to be large enough to span this distance, this implies that the enzyme probably does not move along the DNA. It is possible that both sub-units of the protein may be joined by a flexible hinge as observed for transposon gammadelta resolvase (Abdel-Meguid et al., 1984), thus producing the cleavage of DNA at a variable distance i.e. 25 to 27 bp. Therefore a favourable model could be proposed that the smaller sub-unit of the enzyme binds to the recognition sequence and brings an active site of the restriction sub-unit into contact with an adjacent non-specific DNA sequence.

The interaction between the two domains within the attTn7 could occur by several mechanisms but the most favoured could be, (i) a looping out mechanism, (ii) a directional walk mechanism, or, (iii) binding to a specific site and cutting adjacent to the recognition site (36 bp glmS box) like Type II and Type III restriction enzymes.

The attTn7 site of S. marcescens (1st site) does not fulfil the criteria of looping out (see chapter 1, section 1.7.4). Similarly the experiments described in chapter 7 using

altered spacing clones and the deletion clones also suggest that looping out may not be the case in the attTn7 sequence. Since all altered spacing clones and the deletion clone (Δ 3LOOP) have efficiently promoted the high frequency transposition of Tn7, it is unlikely that the interaction within the attTn7 is brought about by this mechanism. It is likely that Tn7 point of insertion is at a fixed position from the 36 bp glmS box in these clones. However a precise point of insertion of Tn7 in altered spacing clones (MASCPL77, MASCPL73A, MASCPL73B, MASCPL69, MASCPL43 and MASCPL32) and deletion clones (Δ 1LOOP, Δ 2LOOP and Δ 3LOOP) still need to be defined.

In retrospect, a directional walk model can be considered which may require energy (ATP dependent) for the movement of the enzyme (or sub-unit of it) along to the cleavage sites. Since at the insertion site of Tn7 there is the potential to form a cruciform, it may be possible that one domain of the enzyme remains at the site of contact and winds the DNA past it until it comes into contact with one of the possible cleavage sites. Denaturation of DNA at the scission point may not be needed as there is no evidence from this study that the Tn7 insertion point is always located in the potential cruciform structure. The experiments described in chapter 7 have shown the requirement of a cruciform structure in the attTn7 for high frequency transposition of Tn7 but were not sufficient to show the cleavage point within the cruciform structure.

Whether such cruciform structures exist in vivo remain unresolved. However some cruciforms have been shown to occur stably in negatively supercoiled DNA and their formation depends upon temperature, ionic strength, base pair composition and physiological conditions of the cell (Lilley, 1984).

The possible role of a DNA with potential to form cruciform structures adjacent to the 36 bp glmS box in the target replicon could be explained on the following basis.

(i) It may affect the writhing number of Tn7 target replicon thus altering the superhelical density of the molecules. It has been suggested on the basis of several lines of evidence that gene expression and some site-specific recombination might be influenced by DNA supercoiling (Courey and Wang, 1983; Cozzarelli et al., 1984; Lilley, 1984).

(ii) It may allow a flexibility, bending of attTn7 DNA which may be required for efficient high frequency transposition of Tn7. The presence of pseudo-cruciform structure in a linear molecule has been shown to introduce pronounced bending in the DNA (Gough and Lilley, 1985).

(iii) Another possible role may include the recognition of this structure by Tn7 or host encoded proteins. The mechanism of this still remains obscure.

The current dogma indicates that Tn7 transposase binds within the 36 bp glmS box in only one orientation and cleavage by Tn7 transposase occurs at some distance away in one direction.

On the basis of results obtained, it is possible to draw one conclusion that Tn7 transposase may have similarities with the Type II or Type III restriction enzymes. However unlike these enzymes, the Tn7 transposase recognition sequence seems to be much larger and may be composed of at least three separate domains. Other factors that might influence the choice of Tn7 insertion in attTn7 include (i) a potential cruciform at the point of insertion or adjacent to the point of insertion, (ii) DNA bendability or kinks in the attTn7, and (iii) a host factor modulation system.

The Type III enzymes, (for example EcoP15) require ATP (stimulated by S-adenosyle methionine) and cleave the DNA near the recognition site, 25-27 bases in the 3' direction only. There is no apparent indication of sequence specificity at the site of scission in the case of EcoP15 and to a certain extent this is true for Tn7. The similarity between the specificities of Type III and Tn7 transposase is of interest for evolutionary studies.

9.4 Implication of DNA-Protein Interaction:

Since there is no DNA consensus between the ends of Tn7 and the attTn7 sites from different bacterial chromosomes, the question remains: how do the DNA sequences at attTn7 site and at the ends of Tn7 interact with Tn7 transposase(s) in the absence of any homology?

It has been shown that tnsB binds specifically with the 22 bp repeats at the right end of the Tn7 (McKown et al., 1987) and that tnsD binds specifically with the 958 bp attTn7 fragment (David Sherratt, personal communication). Since the two proteins bind separately to the ends of Tn7 and attTn7, it is plausible that the two proteins bound with the respective DNAs interact with each other leading to the site-specific transposition event of Tn7.

Any possible involvement of tnsA or tnsC for the synaptic complex formation by Tn7 ends and attTn7 can also be explained. It may be that the attTn7 site and the ends of Tn7 provide different binding domains to these proteins. Bacteriophage lambda integrase is thought to have multiple affinity to different binding domains at the attP, attB and non-att site DNA (Ross and Landy, 1982). Four well characterised eukaryotic regulators, Gal4, Gcn4, Mat α 2 and Hap1 have been shown to have multiple affinities for unrelated DNA sequences. It was shown in the case of Hap1 that two upstream activation sites are present at Uas1 and Cyc7 UAS. Recent evidence suggests that in both

cases the same protein binds to two sites of unrelated DNA sequence and the precise positions of contact differ in the DNA (Pfeifer et al., 1987).

Whatever molecular mechanism underlies Hap1 recognition of two unrelated sequences could also be responsible for Tn7 and attTn7 sites interactions. This situation may explain how two different attTn7 sites are selected by Tn7 transposase (tnsD) in S. marcescens. This analogy promises to be of great interest in understanding the molecular transpositional recombination reaction and the wide spread of ^rTp in medically important bacteria.

9.5 Future Experiments:

The detailed analysis of the attTn7 site presented in this thesis still needs a greater understanding, (i) of the mechanisms involved in the target site selection by Tn7 encoded proteins, (ii) the site-specific DNA-protein complex formation at the attTn7 site, and (iii) topological studies. Therefore further development is required to facilitate this understanding.

The use of M13 phages as a transducing particles have enabled the definition of the requirement of DNA sequence at attTn7 discussed in the early sections of this chapter. Since the results of the M13 transduction frequency assays have a normal fluctuation of 2 to 5 fold and in the absence of large data sets, it is not possible to calculate any

statistically significant variations in these results. However with this margin of fluctuation, it is still possible to infer a series of observations regarding the properties of attTn7. It is recommended to use some interesting attTn7 sites (i.e. $\Delta 1L00P$, $\Delta 2L00P$ and $\Delta 3L00P$) and to test them using the mating out assay as described by Rogers et al. (1986). However the M13 transduction frequency assay can be improved by using the mini attTn7 site inserted "in frame" in the beta-galactosidase gene. This would allow an easy morphological detection of Tn7 transposition events in the attTn7 site, thus the vectors will fail to produce blue plaques in the presence of IPTG and X-gal.

The 36 bp glmS box discussed in this thesis can be defined accurately from the left end by generating the fine deletion mutants from the left end of this box. Since the transcriptional terminator in the attTn7 has been deleted by generating the ClaI RE site only 5 bp outside this box, it will be easier to establish these deletions.

In order to determine the frequency of transposition of Tn7 in various attTn7 sites, it is necessary to obtain the virgin attTn7 sites from the chromosome of A. tumefaciens, K. pneumoniae and S. marcescens. In an isogenic bacterial strain these attTn7 could be tested for the transposition frequency into them. This would also allow the definition of the Tn7 target duplication sequence in K. pneumoniae

attTn7 site, thereby clarifying the involvement of a ruler in target site selection. Similarly Tn7 target duplication sequence in the altered spacing clones (MASCPL77 and derivatives) and the three deletion clones (Δ 1LOOP, Δ 2LOOP and Δ 3LOOP) would provide conclusive evidence about the spacing between the point of insertion of Tn7 and the last nucleotide of 36 bp glmS box in attTn7. This could be achieved by cloning the various attTn7 sites into a suitable plasmid vector which can be used as the target for Tn7 transposition. The right junction fragments from the recombinant plasmid attTn7::Tn7 can be sub-cloned into M13 phage vectors for DNA sequence determination of the exact Tn7 target duplication sequence in these clones.

The inter-species variation in the distance between the glmS box and the point of insertion (i.e. the additional 11 bp spacing in the non-coding region of the K. pneumoniae attTn7 site) needs to be dissected. Whether it is specified by host factor(s) or that the base composition of K. pneumoniae attTn7 DNA causes this divergence is unclear. This can be answered simply by testing the virgin K. pneumoniae attTn7 site in other host bacterial strains, i.e. E. coli or S. marcescens. The DNA sequence analysis of Tn7-attTn7 junction fragments (produced by Tn7 transposition) would allow the elucidation of clues of this.

Therefore it is important to clone the virgin attTn7 site from the chromosome of K. pneumoniae into a suitable vector

system. In addition, a well-defined in vitro transposition system would provide fundamental clues about the interaction of proteins and DNA at the target site. The highly specific attTn7 sequence defined in this study can be used as a suitable target to probe the steps involved in the transposition of Tn7.

Involvement of an inverted repeat as a recognisable structural feature in supercoiled target molecules can be elucidated by electronmicroscopic studies of attTn7 sites (with and without inverted repeat) incubated with various overproduced tns proteins. An understanding of the involvement of secondary structures in the attTn7 site can be studied by measuring the effect of different known secondary structures (cloned adjacent to the 36 bp glmS box) on the transposition frequency of Tn7.

By using over-expressed Tn7 encoded proteins, it may be easier to define the contact points within the attTn7 DNA by DNase I footprint analysis. Similarly methylation and ethylation protection experiments can provide further clues about the involvement of various domains within attTn7 site.

Within the three defined region of 36 bp glmS box in attTn7, mutations at specific A and T nucleotides can be obtained by using the oligonucleotide directed mutagenesis method. The relative strength of the DNA-protein complexes formed in vitro can be investigated by using the attTn7

sites with different point and multiple mutations.

The DNA sequence analysis of A. tumefaciens attTn7 which could not be completed should be attempted again. Small fragments from the A. tumefaciens::Tn7 chromosome could be cloned in a suitable plasmid vector and identified by using the pAIQ:: Δ Tn7 DNA as a hybridisation probe. It would be very useful to obtain the DNA sequence of the secondary attTn7 sites from the E. coli chromosome. The comparison of these attTn7 sites with the second attTn7 site of S. marcescens may invoke a new pathway for target site selections by Tn7.

REFERENCES

- Abdel-Meguid, S.S., Grindley, N.D.F., Templeton, N.S. and Steitz, T.A. (1984) Cleavage of the site specific recombination protein, gammadelta resolvase: the smaller of the two fragments binds DNA specifically. Proc. Natl. Acad. Sci. USA. 81: 2001-2005.
- Abremski, K., Hoess, R. and Sternberg, N. (1983) Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. Cell 32: 1301-1311.
- Ahmed, A. (1984) Use of transposon-promoted deletions in DNA sequence analysis. J. Mol. Biol. 178: 941-948.
- Akryod, J.E. and Symonds, N. (1983) Evidence for a conservative pathway of transposition of bacteriophage Mu. Nature 303: 84-86.
- Arber, W., Humbelin, M., Caspers, P., Reif, H.J., Iida, S. and Meyer, J. (1980) Spontaneous mutations in the Escherichia coli prophage P1 and IS1-mediated processes. Cold Spring Harbor Symp. Quant. Biol. 45: 38-40.
- Argos, P., Landy, A., Abremski, K., Egan, J.B., Haggard-Ljungquist, E., Hoess, R.H., Kahn, M.L., Kalionis, B., Narayana, S.V., Pierson, L.S., Sternberg III, N. and Leong, J.M. (1986) The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. 5: 433-440.
- Arthur, A. and Sherratt, D.J. (1979) Dissection of the transposition process: a transposon-encoded site-specific recombination system. Mol. Gen. Genet. 175: 267-274.
- Arthur, A., Nimmo, E., Hettle, S. and Sherratt, D. (1984) Transposition and transposition immunity of transposon Tn3 derivatives having different ends. EMBO J. 3: 1723-1729.
- Auerswald, E.A., Ludwig, G. and Schaller, H. (1980) Structural analysis of Tn5. Cold Spring Harbor Symp. Quant. Biol. 45: 107-113.
- Avila, P., de la Cruz, F., Ward, E. and Grinsted, J. (1984) Plasmids containing one inverted repeat of Tn21 can fuse with other plasmids in the presence of Tn21 transposase. Mol. Gen. Genet. 195: 288-293.
- Bachi, B., Reiser, J. and Pirrotta, U. (1979) Methylation and cleavage sequences of the EcoP1 restriction-modification enzyme. J. Mol. Biol. 128: 143-163.

- Backmann, B.J. (1972) Pedigrees of some mutant strains of Escherichia coli K12. Bacteriol. Rev. 30: 525-557.
- Barth, P.T., Datta, N., Hedges, R.W. and Grinter, N.J. (1976) Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistance from R483 to other replicons. J. Bact. 125: 800-810.
- Barth, P.T. and Grinter, N.J. (1977) Map of plasmid RP4 derived by insertion of transposon C TnC. J. Mol. Biol. 113: 455-474.
- Barth, P.T., Grinter, N.J. and Bradley, D.E. (1978) Conjugal transfer of plasmid RP4: Analysis by Transposon 7 insertion. J. Bact. 133: 43-52.
- Bender, J. and Kleckner, N. (1986) Genetic evidence that Tn10 transposes by a nonreplicative mechanism. Cell 45: 801-815.
- Bengtke, J. and Normark, S. (1983) Insertion of IS2 creates a novel ampc promoter in Escherichia coli. Cell 32: 809-816.
- Bennett, P.M., Grinstead, J., Choi, C.L. and Richmond, M.H. (1978) Characterisation of Tn501: a transposon determining resistance to mercuric ions. Mol. Gen. Genet. 159: 101-106.
- Berg, D.E. (1977) Insertion and excision of the transposable kanamycin resistance determinant Tn5. In, DNA insertion elements, plasmids and episomes (Ed. A.I. Bukhari, J.A. Shapiro and S. Adhya). Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York. pp 205-212.
- Berg, D.E. and Berg, L.M. (1983) The prokaryotic transposable element Tn5. Biotechnology 1: 417-435.
- Berg, D.E., Egner, C. and Lowe, J.B. (1983a) Mechanism of F factor-enhanced excision of transposon Tn5. Gene 22: 1-7.
- Berg, D.E., Schmandt, M. and Lowe, J.B. (1983b) Specificity of transposon Tn5 insertion. Genetics 105: 813-828.
- Berg, D.E., Lodge, T., Sasakawa, C., Nag, D.K., Phadnis, S.H., Weston-Hafer, K. and Carle, G.F. (1984) Transposon Tn5: specific sequence recognition and conservative transposition. Cold Spring Harbor Symp. Quant. Biol. 49: 215-226.
- Bickle, T.A. (1982) In, Nucleases (Ed. S.M. Lin and R.J. Roberts). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. pp 85-108.

- Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7: 1513-1523.
- Blackwell, T., Moore, M., Yancopoulos, G., Suh, H., Lutzker, S., Selsing, E. and Alt, F. (1986) Recombination between immunoglobulin variable region segments is enhanced by transcription. Nature 324: 585-589.
- Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R. (1985) Ty elements transpose through an RNA intermediate. Cell 40: 491-500.
- Bolton, E., Glynn, P. and O'Hara, F. (1984) Site specific transposition of Tn7 into a Rhizobium melioli megaplasmid. Mol. Gen. Genet. 193: 153-157.
- Boocock, M.R., Brown, J.L. and Sherratt, D.J. (1986) Structural and catalytic properties of specific complexes between Tn3 resolvase and the recombination site res. Biochem. Soc. Trans. 14: 214-216.
- Bossi, L. and Ciampi, M.S. (1981) DNA sequence at the sites of three insertions of the transposable element Tn5 in the Histidine operon of Salmonella. Mol. Gen. Genet. 183: 406-408.
- Boyer, H.W. and Roulland-Dussoix, D. (1969) A complementation analysis of the restriction and modification of DNA in E.coli. J. Mol. Biol. 41: 459-472.
- Brevet, J., Faure, F. and Borowski. (1985). Tn7-encoded proteins. Mol. Gen. Genet. 201: 258:264.
- Broach, J.R., Guarascio, V.R. and Jayaram, M. (1982) Recombination within the yeast plasmid 2u circle is site-specific. Cell 29: 227-234.
- Brown, N.L. and Smith, M (1977) Cleavage specificity of the restriction endonuclease isolated from Haemophilus gallinarum (HgaI). Proc. Natl. Acad. Sci. USA. 74: 3213-3216.
- Bukhari, A.I. and Zipser, D. (1972) Random insertion of Mu-1 DNA within a single gene. Nature New Biol. (London) 236: 240-243.
- Bukhari, A.I. (1976) Bacteriophage Mu as a transposition element. Ann. Rev. Genet. 10: 389-412.
- Bukhari, A.I. and Ambrosio, L. (1978) The invertible segment of bacteriophage Mu DNA determines the adsorption properties of Mu particles. Nature 271: 575-577.

- Bukhari, A.I., Shapiro, J. and Adhya, S. (1978) Ed. of DNA insertion elements, plasmids and episomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Bukhari, A.I. (1981) Models of DNA transposition. TIBS 6: 56-60.
- Burchardt, J., Weismann J., Hamilton D.L. and Roberts Y. (1981) Complexes formed between the restriction endonuclease EcoK and hetroduplex DNA. J. Mol. Biol. 153: 425-440.
- Caillaud, F. and Courvalin, P. (1987) Tn1545: Nucleotide sequence of the ends of the conjugative shuttle transposon Tn1545. Mol. Gen. Genet. 209: 110-115.
- Calos, M.P. and Miller, J. (1980) Transposable elements. Cell 20: 579-595.
- Campbell, A., Berg, D.E., Botstein, D., Lederberg, E.M., Novick, R.P., Starlinger, P. and Szybalski, W. (1979) Nomenclature of transposable elements in prokaryotes. Gene 5: 197-206.
- Caruso, M. and Shapiro, J.A. (1982) Interaction of Tn7 and temperate phage FII6L of Pseudomonas aeruginosa. Mol. Gen. Genet. 188: 292-298.
- Ciampi, M.S., Schmid, M.B. and Roth, J.R. (1982) Transposon Tn10 provides a promoter for transcription of adjacent sequences. Proc. Natl. Acad. Sci. USA. 79: 5016-5020.
- Chou, J., Casadaban, M.J., Lemaux, P.G. and Cohen, S.N. (1979) Identification and characterization of a self-regulated repressor of translocation of the Tn3 element. Proc. Natl. Acad. Sci. USA. 76: 4020-4024.
- Chou, J., Lemaux, P.G., Casadaban, M.J. and Cohen, S.N. (1979) Transposition protein of Tn3: identification and characterization of an essential repressor-controlled gene product. Nature 282: 801-806.
- Chou, J. and Roizman, B. (1985) Isomerization of Herpes Simplex Virus 1 genome: identification of the cis-acting and recombination sites within the domain of a sequence. Cell 41: 803-811.
- Clewell, D.B., An, F.Y., White, B.A. and Gawron-Burke, C. (1985) Streptococcus faecalis sex pheromone (cAM373) also produced by Staphylococcus aureus and identification of a conjugative transposon (Tn918). J. Bact. 162: 1212-1220.
- Cohen, S.N. (1976) Transposable genetic elements and plasmid evolution. Nature 263: 731-738.

- Cohen, S.N. and Shapiro, J.A. (1980) Transposable genetic elements. Scientific American 242: 36-45.
- Cornelius, G., Ghosal, D. and Saedler, H. (1978) Tn951: a new transposon carrying a lactose operon. Mol. Gen. Genet. 160: 215-224.
- Courey, A.J. and Wang, J.C (1983) Cruciform formation in a negatively supercoiled DNA may be kinetically forbidden under physiological conditions. Cell 33: 817-829.
- Cozzarelli, N.R., Krasnow, S.P., Gerrard, S.P. and White, J.H. (1984) A topological treatment of recombination and topoisomerases. Cold Spring Harbor Symp. Quant. Biol. 49: 383-400.
- Craig, N. and Nash, H.A. (1984) Three mechanisms of phage lambda-site specific recombination: site specific breakage of DNA by Int topoisomerase. Cell 35: 795-803.
- Craigie, R., Huisman, O. and Mizuuchi, K. (1984) Site-specific recognition of the bacteriophage Mu ends by the Mu A protein. Cell 39: 387-394.
- Craigie, R. and Mizuuchi, K. (1985) Cloning of the A gene of bacteriophage Mu and purification of its product, the Mu transposase J. Biol. Chem. 260: 1832-1835.
- Dalarymple, B., Caspers, P. and Arber, W. (1984) Nucleotide sequence of the prokaryotic mobile genetic element IS30. EMBO J. 3: 2145-2149.
- Datta, N., Nugent, M. and Richards, H. (1980) Transposons encoding trimethoprim or gentamicin resistance in medically important bacteria. Cold Spring Harbor Symp. Quant. Biol. 45: 45-51.
- Davis, M.A., Simons, R.W. and Kleckner, N. (1985) Tn10 protects itself at two levels from fortuitous activation by external promoters. Cell 43: 379-387.
- Deho, G., Chisotti, D., Alano, P., Zangrossi, S., Borello, M.G. and Sironi, G. (1984) Plasmid mode of propagation of the genetic element P4. J. Mol. Biol. 178: 191-207.
- Del Rey, F.J., Donahue, T.F. and Fink, G.R. (1982) Sigma, a repetitive element found adjacent to tRNA genes of yeast. Proc. Acad. Sci. USA. 79: 4138-4142.
- Derbyshire, K.M. and Grindley, N.D.F. (1986) Replicative and conservative transposition in bacteria. Cell 47: 325-327.

- Dodd, H.M. and Bennett, P.M. (1987) The R46 site-specific recombination system is a homologue of the Tn3 and (Tn1000) cointegrate resolution system. J. Gen. Microb. 133: 2031-2039.
- Doring, H.-P. and Starlinger, P. (1986) Molecular genetics of transposable elements in plants. Ann. Rev. Genet. 20: 175-200.
- Early, P. and Hood, L. (1981) Allelic exclusion and nonproductive immunoglobulin gene rearrangements. Cell 24: 1-3.
- Echols, H. (1986) Multiple DNA-protein interactions governing high-precision DNA transactions. Science 233: 1050-1056.
- Edlin, G., Lin, L. and Bitner, R. (1977) Reproductive fitness of P1, P2 and Mu lysogens. J. Virol. 21: 560-564.
- Ely, B. (1982) Transposition of Tn7 occurs at a single site on the Caulobacter crescentus chromosome. J. Bact. 151: 1056-1058.
- Emmons, S.W., Yesner, L., Ruen, K-S. and Katzenberg, D. (1983) Evidence for a transposon in Caenorhabditis elegans. Cell 32: 55-65.
- Engler, T.A. and van Bree, M.P. (1981) The nucleotide sequence and protein coding capability of the transposable element IS5. Gene 14: 155-163.
- Engles, W.R. (1983) The P family of transposable elements in Drosophila. Ann. Rev. Genet. 17: 315-344.
- Fennewald, M.A. and Shapiro, J.A. (1979) Transposition of Tn7 in Pseudomonas aeruginosa and isolation of alk::Tn7 mutations. J. Bact. 139: 264-269.
- Finnegan, D.J. (1985) Transposable elements in eukaryotes. Int. Rev. Cytol. 93: 281-326.
- Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S. (1978) Repeated gene families in Drosophila melanogaster. Cold Spring Harbor Symp. Quant. Biol. 42: 1053-1063.
- Fink, G., Farabough, P., Roeder, G. and Chalett, D. (1980) Transposable elements (Ty) in yeast. Cold Spring Harbor Symp. Quant. Biol. 45: 575-580.
- Flavey, E. and Grindley, N.D.F. (1987) Contacts between gammadelta resolvase and the gammadelta res site. EMBO J. 6: 815-821.

- Fling, M. and Richards, C. (1983) The nucleotide sequence of the trimethoprim-resistant dihydrolate reductase gene harbored by Tn7. Nucl. Acids Res. 11: 5147-5158.
- Fling, M., Kopf, J. and Richards, C. (1985) Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme 3"(9)-O-nucleotidyl-transferase. Nucl. Acids Res. 13: 7095-7106.
- Fried, M.G. and Crothers, D.M. (1981) Equilibria and kinetics of the lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucl. Acids Res. 9: 6505-6525.
- Fried, M.G. and Crothers, D.M. (1983) CAP and RNA polymerase interactions with the lac promoter: binding stoichiometry and long range effect. Nucleic Acid Res. 11: 141-158.
- Friedman, D.I., Olson, E.R., Georgopoulos, C., Tilly, K., Herskowitz, I. and Banuett, F. (1984) Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. Microbiol. Rev. 4: 299-325.
- Furth, M., Blattner, F., McLeester, C. and Dove, W. (1977) Genetic structure of the replication origin of bacteriophage lambda. Science 198: 1046-1051.
- Galansdorff, N., Charlier, D. and Zafarullah, M. (1981) Activation of gene expression by IS2 and IS3. Cold Spring Harbor Symp. Quant. Biol. USA. 45: 153-156.
- Galas, D.J., Calos, M.P. and Miller, J.H. (1980) Sequence analysis of Tn9 insertions in the lac Z gene. J. Mol. Biol. 144: 19-41.
- Galas, D.J. and Chandler, M. (1981) On the molecular mechanism of transposition. Proc. Natl. Acad. Sci. USA. 78: 4858-4862.
- Garfinkel, D.J., Simpson R.B., Ream, L.W., White, F.F., Gordon, M.P. and Nester, E.W. (1981) Genetic analysis of crown gall: fine structure map of the T-DNA by site-directed mutagenesis. Cell 27: 143-153.
- Garfinkel, D.J., Boeke, J.D. and Fink, G.R. (1985) Ty element transposition: reverse transcriptase and virus-like particles. Cell 42: 507-517.
- Garner, M.M. and Revzin, A. (1981) A gel method for quantifying the binding of proteins to specific DNA regions: application to the components of the Escherichia cloi lactose operon regulatory system. Nucl. Acids Res. 9: 3047-3060.

- Gay, N.V., Tybulewicz, V.L.J. and Walker, J.E. (1986) Insertion of transposon Tn7 into the Escherichia coli glmS transcriptional terminator. Biochem. J. 234: 111-117.
- Gawron-Burke, C. and Clewell, D.B. (1982) A transposon in Streptococcus faecalis with fertility properties. Nature 300: 281-283.
- Georgiev, G.P., Kramerov, D.A., Ryskov, A.P., Skryabin, K.G. and Lukanidin, E.M. (1984) Dispersed repetitive sequences in eukaryotic genomes and their biological significance. Cold Spring Harbor Symp. Quant. Biol. 49: 1109-1120.
- Gosti-Testu, F., Norris, V. and Brevet, J. (1983) Restriction map of Tn7. Plasmid 10: 96-99.
- Gough, G.W. and Lilley, M.J. (1985) DNA bending induced by cruciform formation. Nature 313: 154-156.
- Grindley, N.D.F. and Sherratt, D.J. (1978) Sequence analysis of IS1 insertion sites: models for transposition Cold Spring Harbor Symp. Quant. Biol. 43: 1257-1261.
- Grindley, N.D.F., Lauth, M.R., Wells, R.G., Wityk, R.J., Salvo, J.J. and Reed, R.R. (1982) Transposon-mediated site-specific recombination: identification of three binding sites for resolvase at the res sites of gammadelta and Tn3. Cell 30: 19-27.
- Grindley, N.D.F. and Reed, R. (1985) Transpositional recombination in prokaryotes. Ann. Rev. Biochem. 54: 863-896.
- Grunstein, M. and Haggness, D. (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA. 72: 3961-3967.
- Guo, L.-H. and Wu, R. (1983) Exonuclease III: use for DNA sequence analysis and in specific deletions of nucleotides. In, Methods in Enzymology (Ed. R. Wu, L. Grossman and K. Moldave) 100: 60-96.
- Habermann, P.P., Klaer, R., Kuhn, S. and Starlinger, P. (1979) IS4 is found between eleven or twelve base pair duplications. Mol. Gen. Genet. 175: 369-373.
- Hadi, S.M., Bachi, B., Shepherd, J.C.W., Yuan, R., Inelchen, K. and Bickle, T.A. (1979) DNA recognition and cleavage by the EcoP15 restriction endonuclease. J. Mol. Biol. 134: 655-666.
- Hadi, S.M., Bachi, B., Iida, S. and Bickle, T.A. (1983) DNA restriction-modification enzymes of phage P1 and plasmid p15B. J. Mol. Biol. 165: 19-34.

- Hakkaart, M.J.J., Van Der Elzen, P.J.M., Veltkamp, E. and Nijkamp, H.J.J. (1984) Maintenance of multicopy plasmid CloDF13 in Escherichia coli cells: evidence for site-specific recombination at parB. Cell 36: 203-209.
- Halford, S., Jordan, S.L. and Kirkbridge, E.A. (1985) The resolvase protein from the transposon Tn21. Mol. Gen. Genet. 200: 169-175.
- Halling, S.M. and Kleckner, N. (1982) A symmetrical six-base pair target site sequence determines Tn10 insertion specificity. Cell 28: 155-163.
- Harshey, R.M. (1984) Transposition without duplication of infecting bacteriophage Mu DNA. Nature 311: 580-581.
- Harshey, R.M. and Bukhari, A.I. (1981) A mechanism of DNA transposition. Proc. Natl. Acad. Sci. USA. 78: 1090-1094.
- Harshey, R.M. and Bukhari, A.I. (1983) Infecting bacteriophage Mu DNA forms a circular-protein complex. J. Mol. Biol. 167: 427-441.
- Hartl, D.L., Dykhuizen, D.E., Miller, R.D., Green, L. and de Framond, J. (1983) Transposable element IS50 improves growth rate of E. coli cells without transposition. Cell 35: 503-510.
- Hassan, D.M. and Brevet, J. (1983) Absence of cis-acting transposition immunity with Tn7. Plasmid 10: 31-44.
- Hatfull, G.F., Noble, S.N. and Grindley, N.D.F. (1987) The gammadelta resolvase induces an unusual DNA structure at the recombinational cross over point. Cell 49: 103-110.
- Hauer, B. and Shapiro, J. (1984) Control of Tn7 transposition. Mol. Gen. Genet. 194: 149-158.
- Hedges, R.W., Datta, N. and Fleming, M.P. (1972) R factors conferring resistance to trimethoprim but not to sulphonamides. J. Gen. Microbiol. 73: 573-575.
- Hedges, R.W. and Jacob, A.E. (1974) Transposition of ampicillin resistance from RP4 to other replicons. Mol. Gen. Genet. 132: 31-40.
- Heffron, F., Bedinger, P., Champoux, J.J. and Falkow, S. (1977) Deletions affecting the transposition of an antibiotic resistance gene. Proc. Natl. Acad. Sci. USA. 74: 702-706.

- Heffron, F., McCarthy, B.J., Ohtsubo, H. and Ohtsubo, E. (1979) DNA sequence analysis of the transposon Tn3: three genes and three sites involved in transposition of Tn3. Cell 18: 1153-1163.
- Heffron, F. (1983) Tn3 and its relatives. In, Mobile genetic elements (Ed. J.A. Shapiro) Academic Press, New York. pp 223-260.
- Heritage, J. and Bennett, P. (1984) The role of tnpA transposase in transposition immunity. Plasmid 12: 218-221.
- Hermann, R.K. and Shaw, J.E. (1987) The transposable genetic element Tc1 in the nematode Caenorhabditis elegans. TIG 3: 222-225.
- Hernalsteens, J.P., De Greve, M., Van Montagu, M. and Schell, J. (1978) Mutagenesis by insertion of the drug resistance transposition of Tn7 applied to the Ti plasmid of Agrobacterium tumefaciens. Plasmid 1: 218-225.
- Herskowitz, I. and Hagen, D. (1980) The lysis-lysogeny decision of phage lambda: Explicit programming and responsiveness. Ann. Rev. Genet. 14: 399-445.
- Hiestand-Nauer, R. and Iida, S. (1983) Sequence of the site-specific recombinase gene cin and of its substrates serving in the inversion of the C segment of bacteriophage P1. EMBO J. 2: 1733-1740.
- Hill, D.E., Hope, I.A., Macke, J.R. and Strunl, K. (1986) Saturation mutagenesis of the yeast H1S3 regulatory site: requirement for transcriptional induction and for binding by GCN4 activator protein. Science 234: 451-457.
- Hirsch, P.R. (1978) Ph.D. thesis, University of East Anglia.
- Hochschild, A.H. and Ptashne, M. (1986) Cooperative binding of lambda repressors to sites separated by integral turns of DNA helix. Cell 30: 9-18.
- Hodges, D.M. (1983) Transposition of the transposon Tn7. Ph.D thesis, University of Sussex.
- Hoeijmakers, T.H.J., Frasch, A.C.C., Bernards, A., Borst, P. and Cross, G.A.M. (1980) Novel expression -linked copies of the genes for variant surface antigens in trypanosomes. Nature 284: 78-80.
- Hulton, C.S.J. (1987) An investigation into host factors involved in Tn7 transposition. M.Sc. thesis, University of London.

- Hyde, D.R. and Tu, C.-P. D. (1985) tnpM: A novel regulatory gene that enhances Tn21 transposition and suppresses cointegrate resolution. Cell 42: 629-638.
- Iida, S., Marcoli, R. and Bickle, T.A. (1981) Variant insertion element IS1 generates 8-base pair duplications of the target sequence. Nature 294: 374-376.
- Iida, S., Meyer, J., Kennedy, K. and Arber, W. (1982) A site-specific conservative recombination system carried by bacteriophage P1. EMBO J. 1: 1445-1453.
- Iida, S., Meyer, J. and Arber, W. (1983) Prokaryotic IS elements. In, Mobile Genetic Elements (Ed. J.A. Shapiro) Academic Press, New York. pp 159-221.
- Iida, S., Hiestand-Nauer, R. and Arber, W. (1985) Transposable element IS1 intrinsically generates target duplication of variable length. Proc. Natl. Acad. Sci. USA. 82: 839-843.
- Iida, S. and Hiestand-Nauer, R. (1986) Localized conversion at the cross over sequences in the site specific DNA inversion of bacteriophage P1. Cell 45: 71-79.
- Irani, M.H., Orosz, L. and Adhya S. (1983) A control element within a structural gene: The gal operon of Escherichia coli. Cell 32: 783-788.
- Isberg, R.R., Lazaar, A.L. and Syvanen, M. (1982) Regulation of Tn5 by the right repeat protein: control at the level of the transposition reaction. Cell 30: 883-892.
- Isberg, R.R. and Syvanen, M. (1982) DNA gyrase is a host factor required for transposition of Tn5. Cell 30: 9-18.
- Jaurin, B. and Normark, S. (1983) Insertion of IS2 creates a novel ampC promoter in Escherichia coli. Cell 32: 809-816.
- Johnson, R.C., Yin, J.-C.P. and Reznikoff, W.S. (1982) Control of Tn5 transposition in Escherichia coli is mediated by protein from the right repeat. Cell 30: 873-882.
- Johnson, R.C. and Simon, M.I. (1985) Hin-mediated site-specific recombination requires two 26 bp recombination sites and a 60 bp recombinational enhancer. Cell 41: 781-791.
- Jordan, E., Saedler, H. and Starlinger, P. (1968) ⁰ and strong polar mutations in the gal operon are insertions. Mol. Gen. Genet. 102: 353-363.

- Kahmann, R. and Kamp, D. (1979) Nucleotide sequences of the attachment sites of bacteriophage Mu DNA. Nature 280: 247-250.
- Kahmann, R., Rudt, F., Koch, C. and Merteris, G. (1985) G inversion in bacteriophage Mu is stimulated by a site within the invertase gene and a host factor. Cell 41: 771-780.
- Kamp, P., Chow, L.T., Broker, T.R., Kwoh, D., Zipser, D. and Kahmann, R. (1979) Site-specific recombination in phage Mu. Cold Spring Harbor Symp. Quant. Biol. 43: 1159-1167.
- Kamp, D. and Kahmann, R. (1981) Two pathways in bacteriophage Mu transposition? Cold Spring Harbor Symp. Quant. Biol. 45: 329-336.
- Kessler, C. and Holtkle, H.-J. (1986) Specificity of restriction endonucleases and methylases. Gene 47: 1-153.
- Khoury, G. and Gruss, P. (1983) Enhancer elements. Cell 33: 313-314.
- Kitts, D.A., Symington, L.S., Dyson, P. and Sherratt, D.J. (1983) Transposon-encoded site-specific recombination: nature of the Tn3 DNA sequences which constitute the recombination site res. EMBO J. 2: 1055-1060.
- Klaer, R., Kuhn, S., Fritz, H.J., Tillmann, E., Saint-Girons, I., Habermann, P., Pfeifer, D. and Starlinger, P. (1980) Studies on transposition mechanisms and specificity of IS4. Cold Spring Harbor Symp. Quant. Biol. 45: 215-224.
- Klar, A., Strathern, J. and Abraham, J. (1984) Involvement of double-strand chromosomal breaks for mating-type switching in Saccharomyces cerevisiae. Cold Spring Harbor Symp. Quant. Biol. 49: 77-88.
- Kleckner, N. (1981) Transposable elements in prokaryotes. Ann. Rev. Genet. 15: 341-404.
- Kleckner, N. (1983) Transposon Tn10. In, Mobile Genetic Elements. (Ed. J.A. Shapiro) Academic Press, New York. pp 261-298.
- Koo, H.-S., Wu, H.-M. and Crothers, D.M. (1986) DNA bending at adenine.thymine tracts. Nature 320: 501-506.
- Kostriken, R., Morita, C. and Heffron, F. (1981) Transposon Tn3 encodes a site-specific recombination system: identification of the essential sequences, genes, and the actual site of recombination. Proc. Natl. Acad. Sci. USA. 78: 4041-4045.

- Kramer, W., Drusta, V., Jansen, H.-W., Kramer, B., Pflugfeder, M. and Fritz, H.-J. (1984) The gapped duplex DNA approach to oligonucleotide-directed mutation construction. Nucl. Acids Res. 12: 9441-9456.
- Krasnow, M.A. and Cozzarelli, N.R. (1983) Site-specific relaxation and recombination by the Tn3 resolvase: recognition of the DNA path between oriented res sites. Cell 32: 1313-1324.
- Krasnow, M.A., Matzuk, M.M., Dungan, J.M., Benjamin, H.W. and Cozzarelli, N.R. (1983) In, Mechanisms of DNA replication and recombination (Ed. N.R. Cozzarelli). Liss, New York. pp 637-659.
- Krebs, M. and Reznikoff, W.S. (1986) Transcriptional and translation initiation sites of IS50. J. Mol. Biol. 192: 781-791.
- Kroger, M. and Hobom, G. (1982) Structural analysis of insertion sequence IS5. Nature 297: 159-162.
- Kushner, S.R. (1978) An improved method for transformation of Escherichia coli with ColE1-derived plasmids. In, Genetic Engineering (Ed. H.B. Boyer and S. Nicosia) Elsevier/North-Holland, Amsterdam. pp 17-23.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.
- Lam, S. and Roth, J.R. (1986) Structural and functional studies of insertion element IS200. J. Mol. Biol. 187: 157-167.
- Laski, F.A., Rio, D.C. and Rubin, G.M. (1986) Tissue specificity of Drosophila P element - transposition is regulated at the level of mRNA splicing. Cell 44: 7-19.
- Lee, C.H., Bhagwatt, A. and Heffron, F. (1983) Identification of a transposon Tn3 sequence required for transposition immunity. Proc. Natl. Acad. Sci. USA. 80: 6765-6769.
- Lichtenstein, C.P. (1980) The unique insertion site of Tn7 in the E.coli chromosome. Ph.D. thesis, University of Cambridge.
- Lichtenstein, C.P. and Brenner, S. (1981) Site-specific properties of Tn7 transposition into the E.coli chromosome. Mol. Gen. Genet. 183: 380-387.
- Lichtenstein, C.P. and Brenner, S. (1982) Unique insertion site of Tn7 in the E.coli chromosome. Nature 297: 601-603.

- Lichtenstein, C.P. and Draper, J. (1985) Genetic engineering of plants. In, "DNA cloning" (Ed. D.M. Glover) IRL Press, Oxford, UK. pp 67-119.
- Liebart, J.C., Ghelardini, P. and Paolozzi, L. (1982) Conservative integration of bacteriophage Mu DNA in pBR322 plasmid. Proc. Natl. Acad. Sci. USA. 79: 4362-4366.
- Lilley, D. M.J. (1984) DNA sequence, structure and supercoiling. Trans. Biochem. Soc. 12: 127-140
- Lin, L., Bitner, R. and Edlin, G. (1977) Increased reproductive fitness of Escherichia coli lambda lysogens. J. Virol. 21: 554-559.
- Ljungquist, E. and Bukhari, A.I. (1977) State of prophage Mu DNA upon induction. Proc. Natl. Acad. Sci. USA. 74: 3143-3147.
- Machida, C. and Machida, Y. (1987) Base substitutions in transposable element IS1 cause DNA duplication of variable length at the target site for plasmid co-integration. EMBO J. 6: 1799-1803.
- Malamy, M., Fiant, M. and Szybalski, W. (1972) Electron microscopy of polar insertions in the lac operon of Escherichia coli. Mol. Gen. Genet. 119: 207-222.
- Mandel, M. and Higa, A. (1970) Calcium dependent bacteriophage DNA in bacteria. J. Mol. Biol. 53: 154-162.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning A Laboratory Manual. Cold Spring Harbor, Cold Spring Harbor Laboratory, New York.
- Maxwell, A., Craigie, R. and Mizuuchi, K. (1987) B protein of bacteriophage Mu is an ATPase that preferentially stimulates intermolecular DNA strand transfer. Proc. Natl. Acad. Sci. USA. 84: 699-703.
- McClintock, B. (1951) Chromosomal organisation and genic expression. Cold Spring Harbor Symp. Quant. Biol. 16: 13-47.
- McClintock, B. (1955) Controlling elements and the gene. Cold Spring Harbor Symp. Quant. Biol. 21: 197-216.
- McKown, R., Waddell, C., Arciszewska, L. and Craig, N.L. (1987) Identification of a transposon Tn7-dependent DNA binding activity that recognizes the ends of Tn7. Proc. Natl. Acad. Sci. USA. 84: 7807-7811.

- McKnight, S.L., Gavis, E.R., Kingsbury, R. and Axel, R. (1981) Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell 25: 385-398.
- Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P.H. (1977) Filamentous cociphage M13 as a cloning vehicle: insertion of a HindIII fragment of the lac regulatory region in M13 replicative form in vitro. Proc. Natl. Acad. Sci. USA. 74: 3642-3646.
- Messing, J. and Vieira, J. (1982) A new pair of M13 phage vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19: 269-276.
- Messing, J. (1983) New M13 vectors for cloning. Methods in Enzymology (Ed. by R. Wu, L. Grossman and K. Moldave) 100: 60-96.
- Miller, D.W. and Miller, L.K. (1982) A virus mutant with an insertion of a copia-like transposable element. Nature 299: 562-564.
- Miller, J.H. (1972) Experiments in molecular genetics. Cold Spring Harbor, Cold Spring Harbor Laboratory, New York.
- Miller, J.H., Calos, M.P., Galas, D., Hofer., M., Buchel, D.E. and Muller-Hill, B. (1980) Genetic analysis of transposition in the lac region of E.coli. J. Mol. Biol. 144: 1-18.
- Miller, H.I. (1984) Primary structure of the himA gene of Escherichia coli: homology with DNA-binding protein Hu and association with the phenylalanyl-tRNA synthetase operon. Cold Spring Harbor Symp. Quant. Biol. 49: 691-698.
- Mizuuchi, K. (1983) In vitro transposition of bacteriophage Mu: a biochemical approach to a novel replication reaction. Cell 35: 785-794.
- Morgan, E.A. (1986) Antitermination mechanisms of rRNA operons of Escherichia coli. J. Bact. 168: 1-5.
- Morita, M., Tsunasawa, S. and Sugino, Y. (1987) Overproduction and purification of the Tn3 transposase. J. Biochem 101: 1253-1264.
- Moore, R. and Krishnapillai, V. (1982) Tn7 and Tn7 insertions into Pseudomonas aeruginosa plasmid R91-5: mapping of two transfer regions. J. Bact. 149: 276-283.
- Morrison, A. and Cozzarelli, N.R. (1979) Site-specific cleavage of DNA by E. coli DNA gyrase. Cell 17: 175-184.

- Murphy, E. (1983) Inhibition of Tn554 transposition: deletion analysis. Plasmid 10: 260-269.
- Murphy, E. and Lofdahl, S. (1984) Transposition of Tn554 does not generate a target duplication. Nature 307: 292-294.
- Murphy, E., Huwyler, L. and de Freire Bastos, M. (1985) Transposon Tn554: complete nucleotide sequence and isolation of transposition-defective and antibiotic-sensitive mutants. EMBO J. 4: 3357-3365.
- Myers, R.M., Lerman, L.S. and Maniatis, T. (1985) A general method for saturation mutagenesis of cloned DNA fragments. Science 229: 242-246.
- Nash, H. (1981) Integration and excision of bacteriophage lambda: the mechanism of conservative site-specific recombination. A. Rev. Genet. 15: 143-167.
- Nash, J. and Krishnapillai, V. (1987) DNA sequence analysis of host range mutants of the promiscuous IncP-1 plasmids R18 and R68 with insertions in oriV. Plasmid 18: 35-45.
- Nasmyth, K.A. and Tatchell, K. (1980) The structure of transposable yeast mating type loci. Cell 19: 753-764.
- O'Hare, K. and Rubin, G.M. (1983) Structure of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. Cell 34: 25-35.
- Ohtsubo, H. and Ohtsubo, E. (1977) Repeated DNA sequences in plasmids, phages and bacterial chromosomes. In, DNA insertion elements, plasmids and episomes. (Ed. A.I. Bukhari, J.A. Shapiro and S.L. Adhya). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. pp 49-63.
- Ohtsubo, H., Ohmori, H. and Ohtsubo, E. (1979) Nucleotide-sequence analysis of Tn3 (Ap): Implications for insertion and deletion. Cold Spring Harbor Symp. Quant. Biol. 43: 1269-1277.
- Ohtsubo, E., Ohtsubo, H., Doroszkiewicz, W., Nyman, K., Alllen, D and Davison, D. (1984) An evolutionary analysis of iso-IS1 elements from Escherichia coli and Shigella strains. J. Gen. App. Microbiol. 30: 359-376.
- Ouartsi, A., Borowski, D. and Brevet, J. (1985) Genetic analysis of Tn7 transposition. Mol. Gen. Genet. 198: 221-227.
- Pfeifer, K., Prezant, T. and Guarente, L. (1987) Yeast HAP 1 activator binds to two upstream activation sites of different sequence. Cell 49: 19-27.

- Philips, S. and Novick, R.P. (1979) Tn554 - A site-specific repressor-controlled transposon in Staphylococcus aureus. Nature 278: 476-478.
- Piekarowicz, A., Bickle, T.A., Shepherd, J.C.W. and Inelchen, K. (1981) The DNA sequence recognised by the HinfIII restriction endonuclease. J. Mol. Biol. 146: 167-172.
- Pierson III, L.S. and Kahn, M.L. (1987) Integration of satellite bacteriophage P4 in Escherichia coli. J. Mol. Biol. 196: 487-496.
- Pilacinski, W., Mosharrafa, E., Edundson, R., Zissler, J., Fiant, M. and Sybalski, W. (1977) Insertion sequence IS2 associated with int constitutive mutants of bacteriophage lambda. Gene 2: 61-74.
- Plasterk, R.H.A., Brinkman, A. and van de Putte, P. (1983) DNA inversions in the chromosome of Escherichia coli and in bacteriophage Mu: relationship to other site-specific recombination systems. Proc. Natl. Acad. Sci. USA. 80: 5355-5358.
- Plasterk, R.H.A. and van de Putte, P. (1984) Genetic switches by DNA inversion in prokaryotes. Biochim. Biophys. Acta 782: 111-119.
- Plasterk, R.H.A. and van de Putte, P. (1985) The invertible P-DNA segment in the chromosome of Escherichia coli. EMBO J. 4: 237-242.
- Ptashne, M. (1986) Gene regulation by proteins acting nearby and at a distance. Nature 322: 697-701.
- Ray, D.S. (1978) In vivo replication of filamentous phage DNA. In, "The single-stranded DNA phages" (Ed. D.T.Denhart, D.Dressler and D.S. Ray) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. pp 325-339.
- Read, H.A., Sarma, S.D. and Jaskunas, S.R. (1980) Fate of donor insertion sequence IS1 during transposition. Proc. Natl. Acad. Sci. USA. 77: 2514-2518.
- Reed, R.R. (1981) Transposon-mediated site-specific recombination: a defined in vitro system. Cell 25: 713-719.
- Reed, R.R. and Grindley, N.D.F. (1981) Transposon-mediated site-specific recombination in vitro: DNA cleavage and protein-DNA linkage at the recombination site. Cell 25: 721-728.

- Reed, R.R. and Moser, C.D. (1984) Resolvase-mediated recombination intermediates contain a serine residue covalently linked to DNA. Cold Spring Harbor Symp. Quant. Biol. 49: 245-249.
- Reynolds, A., Felton, J. and Wright, A. (1981) Insertion of DNA activates the cryptic bgl operon in E.coli K 12. Nature 293: 625-629.
- Rigby, P., Dieckman, M., Rhodes, C. and Berg, P. (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase 1. J. Mol. Biol. 113: 237-251.
- Roberts, D., Hoopes, B.C., McClure, W.R. and Kleckner, N. (1985) IS10 transposition is regulated by DNA adenine methylation. Cell 43: 117-130.
- Roeder, P. and Fink, G.R. (1983) Transposable elements in yeast. In, Mobile Genetic Elements (Ed. J.A. Shapiro) Academic Press, New York. pp 299-326.
- Rogers, M., Ekaterinkai, N., Nimmo, E. and Sherratt, D.J. (1986) Analysis of Tn7 transposition. Mol. Gen. Genet. 205: 550-556.
- Rosen, E., Sivasten, A. and Firtel, R.A. (1983) An unusual transposon encoding inducible and developmentally regulated transcripts in Dictyoslelium. called TDD1. Cell 35: 243-251.
- Ross, W. and Landy, A. (1982) Bacteriophage lambda int protein recognizes two classes of sequence in the phage att site: characterization of arm-type sites. Proc. Natl. Acad. Sci. USA. 79: 7724-7728.
- Rothstein, S.J. and Reznikoff, W.S. (1981) The functional differences in the inverted repeats of Tn5 are caused by single basepair nonhomology. Cell 23: 191-195.
- Rubin, G.M., Kidwell, M.G. and Bingham, P.M. (1982) The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell 29: 987-994.
- Saedler, H., Reif, H.J., Hu, S. and Davidson, N. (1974) IS2, a genetic element for turn-off and turn-on of gene activity in E.coli. Mol. Gen. Genet. 132: 265-289.
- Sanger, F. Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.
- Sapienza, C. and Doolittle, W.F. (1980) Genes are things you have whether you want them or not. Cold Spring Harbor Symp. Quant. Biol. 45: 177-181.

- Schmitt, R., Bernhard, E. and Mattes, R. (1979) Characterization of Tn1721, a new transposon containing tetracycline resistance genes capable of amplification. Mol. Gen. Genet. 172: 53-65.
- Schmitt, R., Rogowsky, P., Halford, S.E. and Grinsted, J. (1985) Transposable elements and evolution. In, Evolution of prokaryotes (Ed. K.H. Schleifa and E. Stacherbrandt). Academic Press, London. pp 91-104.
- Schnabel, H.P, Palm, P., Dick, K. and Grampp, B. (1984) Sequence analysis of the insertion element ISH1.8 and of associated structural changes in the genome of phage oH of the archibacterium Halobacterium halobium. EMBO J. 3: 1717-1722.
- Schreier, P.H. and Cortese, R. (1979) A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13. J. Mol. Biol. 129: 169-172.
- Schurter, W. and Holloway, B.W. (1987) Interaction between the transposable element IS21 on R68.45 and Tn7 in Pseudomonas aeruginosa PAO. Plasmid 17: 61-64.
- Shapiro, J.A. (1969) Mutations caused by the insertion of genetic material into the galactose operon of Escherichia coli. J. Mol. Biol. 40: 93-105.
- Shapiro. J.A. (1979) Molecular model of transposition and replication of bacteriophage Mu and other transposable elements. Proc. Natl. Acad. Sci. USA 76: 1933-1937.
- Shapiro, J.A. (1983) Ed. of Mobile Genetic Elements. Academic Press, New York.
- Sherratt, D.J., Arthur, A. and Burke, M. (1980) Transposon specified site specific recombination systems. Cold Spring Harbor Symp. Quant. Biol. 45: 275-281.
- Shortle, D. and Nathans, D. (1978) Local mutagenesis: A method for generating viral mutants with base substitution in preselected regions of the viral genome. Proc. Natl. Acad. Sci. USA 75: 2170-2174.
- Shortle, D. and Bostein, D. (1983) Directed mutagenesis with sodium bisulfite. Methods in Enzymology (Ed. R. Wu, L. Grossman and K. Moldave) 100: 457-468.
- Simons, R.W. and Kleckner, N. (1983) Translational control of IS10 transposition. Cell 34: 683-691.

- Simsek, M., Das-Sharma, S., Raj-Bhandary, U. and Khorana, H.G. (1982) A transposable element from Halobacterium which inactivates the bacteriorhodopsin gene. Proc. Natl. Acad. Sci. USA. 79: 7268-7272.
- Smith, M. (1982) Site-directed mutagenesis. TIBS. 7: 440-442.
- Smith, G.M. and Jones, P. (1984) Effects of deletions in transposon Tn7 on its frequency of transposition. J. Bact. 157: 962-964.
- Smith, G.M. and Jones, P. (1986) Tn7 transposition: a multigene process. Identification of a regulatory gene product. Nucl. Acids Res. 14: 7915-7927.
- Smith, H.O. and Nathens, D. (1973) A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. J. Mol. Biol. 81: 419-423.
- So, M., Heffron, F. and McCarthy, B.J. (1979) The E. coli gene encoding heat stable toxin is a bacterial transposon flanked by inverted repeats of IS1. Nature 277: 453-455.
- Southern, E. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.
- Spradling, A.C. and Rubin, G.M. (1981) Transposition of cloned P elements in Drosophila melanogaster germ line chromosomes. Science 218: 341-347.
- Starlinger and Saedler (1972) IS elements in microorganisms. Biochemie 54: 177-184.
- Strauss, F. and Varskavaky, A. (1984) A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37: 889-901.
- Streck, R.D., MacGaffey, J.E. and Beckendorf, S.K. (1986) The structure of hobo transposable elements and their insertion sites. EMBO J. 5: 3615-3623.
- Summers, D.K. and Sherratt, D.J. (1984) Multimerization of high copy number plasmids cause instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. Cell 36: 1097-1103.
- Surin, B., Jans, D., Fimmel, A., Shaw, D., Cox, G. and Rosenberg, H. (1984) Structural gene for the phosphate-repressible phosphate-binding protein Escherichia coli has its own promoter: complete nucleotide sequence of the phoS gene. J. Bact. 157: 722-778.

- Syvanen, M. (1984) The evolutionary implications of mobile genetic elements. Ann. Rev. Genet. 18: 271-293.
- Takahashi, K., Vigneron, M., Matthes, H., Wilderman, A., Zenke, M. and Chambon, P. (1985) Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. Nature 319: 121-126.
- Taylor, A.L. (1963) Bacteriophage-induced mutation in E. coli. Proc. Natl. Acad. Sci. USA. 50: 1043-1051.
- Temin, H.M. (1980) Origin of retroviruses from cellular moveable genetic elements. Cell 21: 599-600.
- Temin, H.M. (1985) Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons and retrotranscripts. Mol. Biol. Evol. 2: 455-468.
- Thompson, J.F., de Vargas, L.M., Koch, C., Kahmann, R. and Landy, A. (1987) Cellular factors couple recombination with growth phase: characterization of a new component in the lambda site-specific recombination pathway. Cell 50: 901-908.
- Thomsen, B., Mollerup, S., Bonven, B.J., Frank, R., Blockner, H., Nielsen, O.F. and Westergaard, O. (1987) Sequence specificity of DNA topoisomerase I in the presence and absence of camptothecin. EMBO J. 6: 1817-1823.
- Thomson, J.A., Hendson, M. and Magnes, R.M. (1981) Mutagenesis by insertion of drug resistance transposon Tn7 into a Vibrio species. J. Bact. 148: 374-378.
- Timmerman, K.P. and Tu, C.-P.D. (1985) Complete sequence of IS3. Nucl. Acids Res. 13: 2127-2139.
- Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Improved estimation of secondary structures in ribonucleic acids. Nature New Biol. 246: 40-41.
- Toussaint, A., Faellen, M. and Bukhari, A.I. (1977) Mu mediated illegitimate recombination as an integral part of the Mu life cycle. In, DNA insertion elements, plasmids and episomes (Ed. A.I. Bukhari, J.A. Shapiro and S. Adhya). Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York. pp 275-285.
- Toussaint, A. and Resibois, A. (1983) Phage Mu: Transposition as a life-style. In, Mobile Genetic Elements (Ed. J.A. Shapiro) Academic Press, New York. pp 105-158.

- Towner, K.J., Venning, B.M. and Pinn (1982) Occurrence of transposable trimethoprim resistance in clinical isolates of Escherichia coli devoid of self transmissible resistance plasmids. Antimicrob. Agents Chemother. 21: 336-338.
- Travers, A. and Klug, A. (1987) DNA wrapping and writhing. Nature 327: 987-994.
- Tu, C.P. and Cohen, S.N. (1980) Translocation specificity of the Tn3 element: characterization of sites of multiple insertions. Cell 19: 151-160.
- Turner, P., Barber, C. and Daniels, M. (1984) Behaviour of the transposons Tn5 and Tn7 in Xanthomonas campestris pv. campestris. Mol. Gen. Genet. 195: 101-107.
- Tyler, J. and Sherratt, D.J. (1975) Synthesis of E. colicins in E. coli. Mol. Genet. Genet. 140: 349-353.
- Ulanovsky, L.E. and Trifonov, E.N. (1987) Estimation of wedge components of curved DNA. Nature 326: 720-723.
- Varmus, H.E. (1982) Form and function of retroviral proviruses. Science 216: 812-820.
- Volkert, F.C. and Broach, J.R. (1986) Site-specific recombination promotes plasmid amplification in yeast. Cell 46: 541-550.
- Walker, J.E., Gay, N.J., Saraste, M. and Emberle, A.N. (1984) DNA sequence around the Escherichia coli unc operon: completion of the sequence of a 17 kb segment containing asnA, oriC, unc, glmS and phoS. Biochem. J. 224: 799-815.
- Wallace, L.J., Ward, J.M., Bennett, P.M., Robinson, M.K. and Richmond, M.H. (1980) Transposition immunity. Cold Spring Harbor Symp. Quant. Biol. 45: 183-188.
- Wu, H.-M. and Crothers, D.M. (1984) The locus of sequence directed and protein-induced DNA bending. Nature 308: 509-513.
- Zakour, R.A. and Loeb, L.A. (1982) Site-specific mutagenesis by error-directed DNA synthesis. Nature 295: 708-710.
- Zieg, J., Silverman, H., Hilman, M. and Simon, M. (1977) Recombinational switching for gene expression. Science 196: 170-172.
- Zieg, J. and Simon, M. (1980) Analysis of the nucleotide sequence of an invertible controlling element. Proc. Natl. Acad. Sci. USA. 77: 4196-4200.

-Zoller, M.J. and Smith, M. (1983) Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single stranded DNA template. In, Methods in Enzymology (Ed. R. Wu, L. Grossman and K. Moldave). 100: 468-500.