Cre-loxP mediated genomic targeting to develop rapid and reproducible expression of recombinant proteins in mammalian cells.

Bernard Kwabi-Addo (M.Sc.)

A thesis presented for the degree of Doctor of Philosophy at (QMW)
University of London, 1997.

Abstract

Expression levels of transgenes in mammalian cells show extreme variability between individual clones isolated from a single transfection. This is due to differing number of copies integrated into the genome and also from chromosomal "position effects". Therefore extensive screening is required to isolate a suitable cell line for high level expression of a recombinant protein.

In this work, the Cre-loxP site-specific recombination system was investigated to eliminate such problems by directing the rapid targeting of any input DNA to a single pre-selected site in Chinese hamster ovary (CHO) cell line. Cre is a 38 kDa recombinase encoded by the bacteriophage P1 which mediates recombination between a pair of specific 34 bp target sequences called loxP sites.

The recombination reaction was first investigated *in vitro* to establish which kinetic parameters could be relevant for efficient gene targeting: a recombinant baculovirus was constructed with a hexa-histidine-*cre* fusion gene. The activity of Cre protein purified (by single step, hexa-histidine/nickel-binding, affinity chromatography) from infected insect cells was verified by: (i) Cre-loxP interaction in gel retardation assays and (ii) Cre-mediated intramolecular excision between two loxP sites flanking a *lacZ* gene in a plasmid DNA.

To investigate Cre-mediated targeting of an exogenous DNA to a chromosomal loxP site, three CHO cell lines were constructed, each carrying a loxP site between a β-actin promoter and a secreted alkaline phosphatase (SAP) gene as a reporter. To demonstrate the targeting event, a promoterless *lacZ*/neo^r gene was co-transfected with either a *cre* plasmid or the purified Cre protein from the baculovirus/insect system. Proper targeting should activate expression of β-galactosidase from the chromosomal β-actin promoter and give loss or reduction of SAP expression in G418 resistant transformants. Southern blot analysis showed targeted events mediated by both *cre* plasmid and recombinant Cre protein.

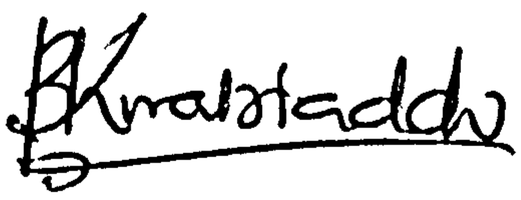
This work should allow the development of a generic mammalian cell line by incorporating the Cre-loxP system for rapid and reproducible large scale production of recombinant proteins.

Acknowledgements

First, I would like to thank the Medical Research Council and The Wellcome Foundation Limited for the award of a post-graduate CASE studentship.
I would like to thank my supervisors, Professor Conrad Lichtenstein at Queen Mary and Westfield College, London, Dr. Tim Peakman at Glaxo Wellcome Medicines Centre and Dr Dirk Gewert at Gemini Research Labs for their help and advice during the project.
In addition, I would like to thank Dr. Nigel Sharp (GWR&D) for his great help and support.
There are many people to thank for their help; Conny Rossmann and Micheal Jowett for tissue culture; Caroline Griffiths and Jenny Worden for baculovirus work, Carolyn Dent and Jon Ellis for their advice on DNA-protein interaction studies. Jane Richardson for useful discussion on the Cre-loxP work and Neil Clark for general advice. Also Andrew Smith at Edinburgh for advice and help.
Lastly, thanks to Fela, Dale and Sophie for their encouragement and to my aunt for her great support.

Declaration

I hereby declare that this thesis has been composed by myself and has not been accepted in any previous application for a degree. The work presented has been carried out by myself unless otherwise stated. All sources of information have been acknowledged by means of reference and any quotations have been distinguished by quotation marks.



Bernard Kwabi-Addo.

Table of Contents

	page
Abstract	2
Acknowledgements	3
Table of contents	4
List of Figures	11
List of Tables	14
Abbreviations	15
CHAPTER 1: Introduction	17
1.1 Mammalian expression systems	19
1.2 Strategies to increase expression levels in mammalian cells	20
1.2.1 Promoter strength, transcriptional control elements	21
1.2.2 Increasing gene copy number: use of viral replicons	22
1.2.2.1 Viral replicons	22
1.2.2.2 Selection for the integration of multiple copy following transfection	23
1.2.2.3 Use of an amplifiable, selectabe marker gene	24
1.2.2.4 *In vitro* amplification	26
1.2.3 Overcoming position effects	27
1.2.3.1 Transfection with a locus control region	27
1.2.3.2 Targeted integration using site-specific recombination system	28
1.3 The Bacteriophage P1 Cre-loxP site-specific recombination system	29
1.3.1 The role of Cre-loxP recombination system in the life cycle of bacteriophage P1	30
1.3.1.1 Role of Cre-loxP recombination in the cyclisation of P1 DNA	30
1.3.1.2 Role of Cre-loxP in P1 vegetative growth and lysogeny in host cell	31
1.3.1.3 Role of Cre-loxP in maintenance of unit-copy P1 plasmid prophage	31
1.3.2 Cre is a member of of the int family of recombinases	32
1.3.3 The Cre recombinase	34
1.3.4 The loxP site	35
1.3.5 Cryptic loxP site(s)	37

1.3.6 The mechanism of Cre-loxP recombination reaction 37
1.3.7 Factors required for the Cre-loxP reaction 39
1.3.8 Topology of DNA substrate required for the recombination reaction 39
1.4 Application of the Cre-loxP site-specific recombination system 41
1.4.1 Cre-loxP recombination in plants 42
1.4.2 Cre-loxP mediated gene targeting in mammalian cells 43
1.4.3 Manipulation of transgenes by Cre-loxP recombination system 44
1.4.4 Other applications of the Cre-loxP site-specific recombination system 48
1.5 Disadvantage(s) of the Cre-loxP recombination system 50
1.6 Aims of the project 53

CHAPTER 2: Materials and Methods 54
2.1 Materials 54
2.2 General DNA subcloning techniques 54
2.2.1 Restriction endonuclease digestion of DNA 54
2.2.2 Ethanol precipitation of DNA from solution 54
2.2.3 Extraction of a DNA solution with organic solvent 55
2.2.4 Agarose gel electrophoresis of DNA 55
2.2.5 Recovery of DNA from an agarose gel 55
2.2.6 Conversion of DNA fragments with 5' overhangs to blunt-ended fragments using klenow 55
2.2.7 Conversion of DNA fragments with 3' overhangs to blunt-ended fragments using T4 DNA polymerase 56
2.2.8 SI nuclease digestion 56
2.2.9 Dephosphorylation of linear DNA fragments using calf intestinal alkaline phosphatase 56
2.2.10 Phosphorylation of linear DNA fragments using bacteriophage T4 polymerase kinase 57
2.2.11 Ligation of DNA fragments 57
2.2.12 Preparation of *E. coli* DH5α competent cells 57
2.2.13 Transformation of *E.coli* DH5α competent cells 58
2.2.14 Small-scale plasmid preparation 58
2.2.15 Large-scale plasmid preparation 59
2.216 Genomic DNA preparation 60
2.2.17 Determination of DNA concentration 60

2.2.18 Southern blotting 61
2.2.19 Preparation of radioactively labelled DNA probes 61
2.2.20 Hybridisation of Southern blot 62
2.2.21 DNA sequencing 62
2.2.22 General polymerase chain reaction (PCR) 63
2.3 Mammalian cell culture techniques 64
2.3.1 Cell line 64
2.3.2 Cell culture media 64
2.3.2.1 Complete medium 64
2.3.2.2 Select medium 64
2.3.3 Maintenance of CHO-DUKX B11 cell line 64
2.3.4 Freezing cell lines for long term storage in liquid nitrogen 65
2.3.5 Reviving frozen cells 65
2.4 Mammalian cell transfection techniques 66
2.4.1 Transient transfection of CHO-DUKX B11 cell using Transfectam 66
2.4.2 Stable transfection of CHO-DUKX B11 cells using Transfectam 66
2.4.3 Dilution cloning of stably transfected CHO-DUKX B11 cells 66
2.4.4 Targeted transfection of CHO-DUKX B11 transformants using LipofectACE 67
2.4.5 Stable transfection of CHO-DUKX B11 cells by electroporation 67
2.4.6 Expanding CHO-DUKX B11 transformed clones 68
2.5 Insect cell and Baculovirus Techniques 68
2.5.1 Maintenance of Sf9 cell line 68
2.5.2 Production of recombinant baculovirus 69
2.5.3 Plaque purification of recombinant baculovirus 69
2.5.4 Preparation of a high titrevirus stock 70
2.5.5 Titration of recombinant virus 70
2.5.6 Infection of Sf9 cells 71
2.5.7 Recombinant baculovirus DNA preparation 71
2.5.8 Screening recombinant virus by the polymerase chain reaction (PCR) 72
2.6 Protein Techniques 72
2.6.1 Polyacrylamide gel electrophoresis (PAGE) of protein samples 72

2.6.2 Metabolic labelling of proteins produced by the recombinant baculoviral vector 73
2.6.3 Purification of Cre recombinase from immobilised Ni^{++}-spin column 73
2.6.4 Determination of protein concentration 75
2.7 Protein-DNA interactions in band shift assays 75
2.7.1 Non-denaturing polyacrylamide gels 75
2.7.2 Band shift assays 75
2.7.2.1 Synthesis of oligonucleotides 75
2.7.2.1b Annealing of oligonucleotide sequence 76
2.7.2.2 5' labelling of oligonucleotide with ^{32}P 76
2.7.2.3 Removal of unincorporated γ-^{32}P dCTP 76
2.7.2.4 DNA-protein interaction in band shift assay 77
2.7.3 "On rate-off rate" band shift competition assays 77
2.8 *In vitro* assay systems 78
2.8.1 *In vitro* Cre-mediated recombination assay 78
2.8.1.1 Restriction digestion and Southern hybridisation analysis 78
2.8.1.2 Restriction digestion and quantification by Fluorescence 78
2.8.1.3 Analysis of recombinant event by phenotypic blue/white colour selection 79
2.8.1.3b Analysing recombinants by sequencing 79
2.8.2 CAT ELISA assay of transformed CHO-DUKX B11 cells 79
2.8.3 Secreted Alkaline Phosphatase (SAP) assay 80
2.8.4 Determination of Specific Production Rates (SPR) 80
2.8.5 β-gal staining of transformed CHO-DUKX B11 clones 81
2.8.6 β-galactosidase enzyme assay of transformed CHO-DUKX B11 81

CHAPTER 3: Expression and purification of bacteriophage P1 Cre recombinase in insect cells using a recombinant baculovirus vector 83
3.1 Introduction 83
3.2 The biology of baculoviruses 84
3.2.1 The genetic organization of baculoviruses 84
3.2.2 Mechanism of gene regultation 86
3.3 Process of infection in insects 86

3.4 Use of baculovirus infected insect cells for the expression of heterologous proteins 89
3.4.1 Development of recombinant baculovirus expression vectors 89
3.4.1.1 Transfer vectors 90
3.4.1.2 Advances in the isolation of recombinant baculovirus 90
3.4.2 Recombinant protein expression 93
3.4.3 Limitations to the baculovirus/insect cell expression system 94
3.4.3.1 Expression levels 94
3.4.3.2 Quality of recombinant protein 95
3.4.3.3 Overcoming limitations by the use of earlier viral promoters 96
3.5 Results of Cre recombinase expression and purification in the baculovirus/insect system 98
3.5.1 Construction of chimeric transfer vector pAcH6N1-Cre 98
3.5.2 Production of recombinant baculovirus containing *cre* gene 101
3.5.3 Metabolic labelling of infected cells 106
3.5.4 Purification of Cre protein from immobilised Ni^{++}-NTA spin column 108
3.6 Construction of pAT-lacZloxP2 110
3.7 Analysing the biological activity of.the.recombinant.Cre.protein 111
3.7.1 Restriction endonuclease digestion analysis of Cre-mediated recombination 113
3.7.2 Analysing the results of Cre-mediated recombination *in vivo* 113
3.7.3 Sequence analysis of excision products 116
3.8 Determination of the specific activity of the Cre protein 118
3.9 Summary 120

CHAPTER 4: Investigating the mechanism of Cre-mediated recombination *in vitro* 121
4.1 Introduction 121
4.2 Results 124
4.3 Recognition binding interaction using band shift assay 125
4.3.1 Investigating Cre-loxP interaction 125
4.3.2 Investigating the specificity of the Cre protein for loxP binding site 127
4.3.3 Construction of "mutant" substrates for Cre-mediated site-specific recombination 130

4.3.4 Investigating Cre-mediated site-specific recombination in *mutants A and B* 132
4.3.5 "On-Off" binding assay 133
4.4 The effect of temperature on Cre-loxP reaction 135
4.4.1 Temperature effect on the efficiency of the excision reaction 137
4.5 Investigating other requirements for the Cre-mediated intramolecular recombination 139
4.6 Investigating the time course for the reaction 141
4.7 Further investigating of Cre-mediated excision reaction 144
4.8 Intramolecular recombination between two loxP sites in inverted repeats 146
4.9 *In vitro* analysis of integration event 150
4.10 Summary 152

CHAPTER 5: Targeting of exogenous DNA into CHO chromosome mediated Cre-loxP system 154
5.1 Introduction 154
5.1.1 Choice of cell line 155
5.1.2 Methods for introducing transgenes into mammalian cells 155
5.1.3 Design of expression vectors 156
5.2 Design of Cre-loxP mediated targeted integration of transgenes in CHO cells 158
5.3 Results 159
5.4 Construction of pSAP-RDN/loxP 159
5.5 Construction of CHO cell lines containing a single loxP site by lipofection 164
5.5.1 Assay for SAP activity in DHFR+ clones 166
5.5.2 Analysis of loxP copy number by Southern hybridisation 169
5.6 Transfection by electroporation 171
5.6.1 Construction of CHO cell lines containing a single loxP plasmid by electroporation 172
5.6.1.1 Investigating DHFR+ clones by polymerase chain reaction (PCR) 174
5.6.1.2 Analysis of loxP copy number by Southern hybridisation 176
5.6.2 Further transfection by electroporation to isolate CHO cell lines containing a single loxP plasmid 178
5.6.2.1 Investigating SPR of SAP in DHFR+ clones containing a single copy *insertion of loxL plasmid* 182

5.6.3 Summary of CHO transfected cells containing a single copyof loxP site 182
5.7 Transient transfection assays 184
5.7.1 Investigating the activity of Cre expression plasmid in transient transfection assays 185
5.8 Site-specific targeting of transgenes to a chromosomal loxP site 190
5.8.1 Construction of loxP targeting plasmids 190
5.8.2 Cre expression plasmid mediated site-specific targeting of pMAMneoCAT-loxP into cell line 1-G12 194
5.8.3 Cre expression plasmid mediated site-specific targeting of pHM3-loxP into cell line 1-G12 196
5.8.3.1 β-gal staining of G418 resistant clones 197
5.8.3.2 β-gal and SAP assays of targeted clones 198
5.8.3.3 PCR analysis of targeted clones 200
5.8.3.4 Southern blot analysis of targeted clones 200
5.9 Use of recombinant Cre protein in site-specific recombination 207
5.9.1 Cre protein mediated targeted integration into cell line 13-G6 using electroporation 208
5.9.2 Cre protein mediated targeted integration into cell line 13-G6 using lipofectACE 209
5.9.2.1 Southern hybridisation of the targeted event 210
5.9.2.2 Further Southern hybridisation to analyse the targeting event 212
5.9.2.3 Assay for β-gal and SAP expression from targeted clones 215
5.10 Summary 217

CHAPTER 6: Discussion 219
Future Work 229
Appendices 233
References 243

List of Figures

		page
Figure 1.1	Schematic representation of Cre-mediated recombination between two loxP sites	33
Figure 1.2	Sequence of loxP protected by Cre recombinase in bacteriophage P1 in DNA-protein footprinting experiments	36
Figure 1.3	Schematic representation of Cre-mediated targeting of an exogenous DNA to a chromosomal loxP site	52
Figure 3.1	The natural cycle of infection by baculoviruses in insect larvae	87
Figure 3.2	Simplified restriction map of BacPAK6	92
Figure 3.3a	Plasmid map of pAcH6N1	100
Figure 3.3b	Schematic representation of PCR designed to amplify *cre* gene from plasmid pMC1-Cre	102
Figure 3.3c	Agarose gel electrophoresis analysis of PCR product	102
Figure 3.3d	Map of the chimeric transfer vector pAcH6N1-Cre	103
Figure 3.4	Agarose gel separation of PCR reactions of recombinant baculoviral genomic DNA	105
Figure 3.5	The rate of Cre synthesis analysed on an SDS-PAGE	107
Figure 3.6	SDS-PAGE analysis of Cre purification from immobilised Ni^{++}-NTA spin column	109
Figure 3.7	Map of plasmid pAT-lacZloxP2	112
Figure 3.8	Diagrammatic presentation of Cre-mediated excision reaction	114
Figure 3.9	Restriction digestion analysis of the Cre-mediated reaction	115
Figure 3.10	Investigating the recombination product by sequence analysis	117
Figure 3.11	Determining the specific activity of the recombinant Cre protein	119
Figure 4.1	Investigating Cre-loxP interaction in band shift assay	126
Figure 4.2	Investigating the specificity of the Cre protein for loxP binding site	129
Figure 4.3	Schematic representation of intramolecular recombination sites	131
Figure 4.4	"On-Off" binding assay	134
Figure 4.5	Investigating the effect of incubation temperature on the Cre-loxP interaction	136
Figure 4.6	Investigating the effect of temperature on the efficiency of the excision reaction	138
Figure 4.7a	Investigating the time course for the Cre-loxP reaction	142

Figure 4.7b	Quantitative analysis of the Cre-loxP reaction	143
Figure 4.8a	Cre-mediated recombination in plasmid pRH43	145
Figure 4.8b	Investigating the efficiency of the Cre-mediated intramolecular excision event	147
Figure 4.9	Investigating Cre-mediated recombination between loxP sites in inverted repeats	149
Figure 4.10	Investigating Cre-mediated integration event	151
Figure 5.1	Schematic representation of Cre-mediated targeting of an exogenous DNA to a chromosomal loxP site	160
Figure 5.2	Map of pSAP-RDN	162
Figure 5.3a	Map of loxR	163
Figure 5.3b	Map of loxL	165
Figure 5.4	Detection by Southern analysis of pSAP-RDN/loxP transfection into CHO-DUKX B11 cells by transfectam	170
Figure 5.5	Agarose gel separation of PCR reactions of independent pSAP-RDN/loxP transfected CHO-DUKX B11 cells	175
Figure 5.6	Detection of pSAP-RDN/loxP (10 μg) transfection into CHO-DUKX B11 cells using electroporation	177
Figure 5.7a	Detection of pSAP-RDN/loxP (1 μg linearised) transfection into CHO-DUKX B11 cells by electroporation	179
Figure 5.7b	Detection of transformed CHO-DUKX B11 cells containing either a single loxP plasmid or tandem repeats of the loxP plasmid by Southern blot analysis	180
Figure 5.8	Detection of pSAP-RDN/loxP (1 μg linearised) transfection into CHO-DUKX B11 cells by electroporation	181
Figure 5.9	Comparison of the specific production rate (SPR) of SAP for three cell lines each containing a single copy of loxP site	183
Figure 5.10	Map of pMAMneoCAT-loxP	191
Figure 5.11	Map of pHM3-loxP	193
Figure 5.12	Schematic representation of Cre-mediated targeting of an exogenous DNA to a chromosomal loxP site	195
Figure 5.13	Agarose gel separation of PCR reaction of independent clones resulting from pHM3-loxP transfection into cell line 1-G12	201
Figure 5.14a	Linearised map of the chromosomal target before Cre-mediated integration (parental) and after targeting with pHM3-loxP	203
Figure 5.14b	Southern blot analysis of Cre-mediated transfection of pHM3-loxP into cell line 1-G12	204

Figure 5.15 Detection by Southern blot analysis of single-copy integration of pHM3-loxP into cell line 1-G12 206

Figure 5.16 Detection of recombinant Cre protein mediated integration of pHM3-loxP into cell line 13-G6 211

Figure 5.17a Linearised map of the chromosomal target before Cre-mediated integration (parental) and after targeting with pHM3-loxP 213

Figure 5.17b Detection of single-copy integrants corresponding to recombinant Cre protein mediated targeted integration of pHM3-loxP into the parental cell line 13-G6 214

List of Tables

		page
Table 4.1	The ionic requirement for Cre-mediated intramolecular recombination	140
Table 5.1	Comparison of standard production rate (SPR) of SAP expression from independent CHO-DUKX B11 cells transfected with recombinant pSAP-RDN/loxP	168
Table 5.2	Investigating Cre recombinase activity in crude cell extract	187
Table 5.3	Investigating the minimum ratio of substrate to Cre protein required for the *in vivo* site-specific recombination	188
Table 5.4	Estimation of β-galactosidase and SAP (SPR) activity from independent pHM3-loxP transfected 1-G12 cells	199
Table 5.5	Estimation of β-galactosidase and SAP (SPR) activity from independent pHM3-loxP transfected 13-G6 cells	216

Abbreviations

AcNPV	Autographa californica nuclear polyhedrosis virus
Amp^R	β-lactamase resistance gene
ATP	Adenosine triphosphate
β-gal	β-galactosidase
bp	base pair
BSA	Bovine serum albumin
CHO	Chinese hamster ovary cell line
Ci	Curie
Cre	Cause of recombination
DHFR	dihydrofolate reductase
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease I
dNTP	deoxynucleoside triphosphate
ds	double stranded
EDTA	ethylenediamine tetraacetic acid
GWR&D	Glaxo Wellcome Research and Development
IPTG	isopropyl-beta-D-thiogalactopyranoside
kb	kilobase pair
kDa	kilo dalton
loxP	Locus of crossover
NCS	Neocarzinostatin
neo^R	neomycin resistance gene
ori	origin of replication
PCR	polymerase chain reaction
pers. comm.	personal communication
Ph	polyhedrin
pi	post-infection
pfu	plaque forming units
poly A	polyadenylation signal
Probe	Radiolabelled DNA fragment
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	revolution per minute
SAP	secreted alkaline phosphatase
SPR	specific production rate

SV40	simian virus 40
tPA	Tissue plasminogen activator
WRL	Wellcome Research Laboratories
WT	wild type

CHAPTER 1

INTRODUCTION

The advent of recombinant DNA technology in the early 1970s has completely revolutionised cell biology. This technology comprises a mixture of techniques, which includes the specific cleavage of DNA by restriction endonucleases which greatly facilitates the isolation and manipulation of individual genes, the cloning of a specific DNA fragment into a self-replicating genetic element (plasmid or virus vector) which is maintained in bacteria so that a single DNA molecule can be reproduced to generate many billions of identical copies, and also genetic engineering, by which DNA sequences are altered to make modified versions of genes, which are then re-introduced back into cells or organisms.

One of the most important contributions of DNA cloning and genetic engineering to cell biology is that they have made it possible to manufacture any of a cell's proteins in large quantities. Traditionally, proteins were purified from cell extracts, on the assumption that enough of the protein is made in the cell for it to be isolated in the first place. However, many key cellular proteins are made transiently and at very low levels e.g., proteins involved in cell division, lymphokines etc. In principle recombinant DNA technology can be used to produce any protein in large quantities. This is done by first cloning the gene for the protein of interest and then inserting it into a special plasmid called an expression vector. The vector is designed in such a way that when it is introduced into an appropriate type of "expression system" (described below), the inserted gene directs the synthesis of very large amounts of the appropriate protein. The impact of this goes far beyond understanding cellular processes, many proteins of commercial value as pharmaceuticals (e.g., human growth hormone, interferon [IFN-α], interleukin-2) which would otherwise be very expensive to purify are now easily obtained using recombinant DNA technology.

There is a wide variety of biological expression systems which spans bacteria, yeast, plants, baculoviral/insect (described more fully in chapter 3) and mammalian cells for

the expression of recombinant proteins. A summary of the advantages and disadvantages of each host system is described in the table below.

Organism	Advantages	Disadvantages
Escherichia coli	(i) easy to manipulate	(i) does not usually secrete proteins into growth medium
	(ii) promoters and gene regulation well understood	(ii) overexpression of foreign proteins often form aggregates of denatured proteins
	(iii) many high expression vectors are available	(iii) many foreign proteins rapidly degraded
	(iv) easy to culture on large scale (e.g. use in manufacture of insulin, interferon and human somatotrophin).	(iv) many post-translational modifications do not occur.
Saccharomyces cerevisiae	(i) widely used industrial organism which is easy to culture	(i) a lot still to be learned about control of gene expression
	(ii) glycosylates proteins	(ii) post-translational modifications are not necessarily the same as those in animal cell
	(iii) can get export into growth medium of heterologous proteins	(iii) heterologous proteins can form inclusion bodies
	(iv) high-level expression systems developed.	
Baculovirus/Insect	(i) widely used industrial organism which is easy to culture	(i) only suitable for transient expression
	(ii) can export proteins into growth medium	
	(iii) good expression levels	
Mammalian cells	(i) get export of proteins	(i) large-scale growth of animal cells is costly
	(ii) get post-translational modifications and products which are not likely to be immunogenic to humans	(ii) great care is needed to avoid contamination of culture.
	(iii) good expression systems available	

The principal advantage of mammalian cells which has made it the preferred choice for heterologous expression of eukaryotic genes is that signals for synthesis, processing and secretion of higher eukaryotic proteins are properly and efficiently recognised. Consequently, a wide variety of eukaryotic proteins have been successfully expressed which are correctly processed in biologically active form (reviewed in Bebbington and Hentschel, 1985; MacDonald, 1991). Typically when eukaryotic cells are transfected, the input DNA (the transgene) integrates by illegitimate recombination at one or more random target site(s) in the genome. The integration site of the input DNA can have profound effects upon levels of expression of the transgene. Thus, it is generally found that individual stable clones isolated from a single transfection express recombinant proteins at a wide range of levels from a few nanograms to several micrograms per 10^6 cells/day (Kaufman *et al.*, 1985) due to these chromosomal "position effects" and/or differing copy number of transgene (described below). It would be advantageous to overcome these two constraints on production levels, so that individual clones would express recombinant proteins at equivalent and preferably reproducible levels and hence decrease the number of individual clones which must be screened in order to identify a high-expressing clone.

The aim of this project was to use Chinese Hamster Ovary (CHO) cells to investigate targeted integration of transgenes to a pre-selected locus in CHO chromosome, for the reproducible high level expression of recombinant proteins. This chapter reviews mammalian expression systems and the variety of methods used to optimise expression levels. As the work described in this thesis exploited the Cre-loxP site-specific recombination system encoded by bacteriophage P1, the biological and biochemical properties of these are also reviewed as are its application in other systems.

1.1 Mammalian expression systems

The wide variety of transcriptional control elements (reviewed in Levinson, 1990) and selectable markers (Berg *et al.*, 1981; Mulligan and Berg, 1980; Wigler *et al.*, 1977) now available allows the use of essentially any cell line for the purpose of heterologous gene expression.

The key steps in the expression of heterologous genes in mammalian cells includes the introduction of the foreign DNA into cells, selection of cells which have integrated the foreign gene into their genome and their subsequent propagation to generate stable

cell lines. There are two general methods of introducing genetic information into mammalian cells; those mediated by virus infection and those mediated by direct DNA transfer (reviewed in Kaufman, 1990). One advantage of the former relates to its efficiency; it is not unusual to achieve close to targeting 100% of cells by exploiting the ability of the virus to infect cells efficiently. A disadvantage of this approach, and one that limits its utility, lies in the relative difficulty of its application; it is technically more difficult and more time consuming to engineer a virus to express a foreign gene in a manner that does not compromise its function than it does to administer DNA directly. As such, most attempts to express heterologous genes in mammalian cells rely on the construction of DNA vectors and their direct transfer by a number of chemical and physical methods, each with its own advantages and disadvantages (reviewed in Kriegler, 1990). Gage *et al.*, (1992) showed that, at least for the two cell lines and three transfection methods they studied (i.e. established Rat-1 cell lines and rat primary skin fibroblasts; by $CaPO_4$ co-precipitation, lipofection and electroporation) the transfection method does not affect the level of foreign gene expression achieved.

As mentioned above, it is generally found that individual stable clones isolated from a single transfection express the recombinant protein at a wide range of levels due to chromosomal "position effects" and also differing copy number of transgene. Furthermore, the level of protein expression depends on other factors such as the efficiency of transcription, mRNA processing, transport, stability, and translational efficiency and protein processing, secretion and stability. A variety of approaches have been taken to optimise expression levels in mammalian systems which are described in the following section.

1.2 Strategies to increase expression levels in mammalian systems

Much of the work which has been done to improve the yield of recombinant protein from mammalian cells has concentrated on improving the amount of foreign gene transcription, which leads to an increase in product-specific mRNA, and hence to increased recombinant protein synthesis (providing other factors are not limiting). This has been achieved by the use of efficient promoter/enhancer sequences and by increasing the copy number of the foreign gene in the transfected cell. Because transfected DNA essentially integrates randomly in stably transfected cells, efforts have been made to overcome the influence of different chromosomal regions ("position effects") which can result in the low rate of transcription of the integrated

foreign gene. More recently, strategies to improve recombinant protein yield have focused on processes downstream of transcription, for example protein secretion. However, the key features for the control of heterologous gene expression consists of a promoter for transcription initiation in conjunction with an enhancer (a cis-acting DNA sequence which improves gene transcription) and a polyadenylation signal which directs the cleavage and polyadenylation of the transcript. In addition, a selectable marker is included to allow for the selection of stable cell lines.

1.2.1 Promoter strength, transcriptional control elements

The promoter/enhancer elements commonly used to control foreign gene expression may be of viral origin, for example the human cytomegalovirus (hCMV) major immediate early (MIE) promoter/enhancer (Foecking and Hofstetter, 1986) or the Rous sarcoma virus long terminal repeat ([RSV-LTR]; Gorman *et al.*, 1982), or from a constitutively highly expressed non-tissue-specific mammalian cellular gene, for example the human β-actin promoter. Promoters show different activity in different cell types, whilst enhancers may have differential activity with different promoters and/or in different cell types. For this reason, expression vectors are generally constructed to contain a promoter/enhancer combination that is highly active in a range of cell types. Wenger *et al.*, (1994) evaluated the activity of ten different promoter/enhancer combinations after transient transfection of ten commonly used mammalian cell lines and found that they were not equally strong in all cell types, but some e.g. herpes simplex virus (HSV) thymidine kinase (TK) promoter/SV40 early enhancer were reasonably active in all the cell types tested.

Transcriptional control in host systems is dependent on endogenous cellular transcription factors and RNA polymerase, either of which may be limiting. Other systems have been described in which the foreign gene is placed under the control of the bacteriophage T7 promoter and co-expressed in the cell with T7 RNA polymerase, this polymerase being highly specific for the T7 promoter. The resulting mRNA transcripts are not efficiently capped by the addition of a 7-methyl guanosine residue at the 5' end of the translational initiation codon, AUG and hence translation occurs at low efficiency. Inclusion of the encephalomyocarditis virus (EMCV) internal ribosomal entry site (IRES) downstream of the T7 promoter, so that it is at the 5' end of the mRNA, can circumvent this problem. Elroy-Stein *et al.*, (1989) reported 5-10 fold higher expression levels after inclusion of this EMCV region.

Hippenmeyer and Highkin (1993) describe the use of an engineered Baby Hamster Kidney cell line, BHK-21, stably expressing the herpes simplex (HSV)-1 transactivator VP16, which increases transcription from immediate early (HSV)-1 promoters. Two reporter proteins, human tissue plasminogen activator (tPA) and bovine growth hormone, were expressed in the BHK/VP16 cell line, under the control of two different immediate early promoters. Stably transfected clones were isolated that expressed protein at higher levels compared to control BHK-21 transfectants (for example 1-15 μg tPA/10^6 cells/day compared to under 0.2 μg/10^6 cells/day). In addition, the BHK/VP16 system compared favourably with two other strategies commonly used to achieve high levels of expression in mammalian cells: the standard Chinese Hamster Ovary cell/dihydrofolate reductase co-amplification system (discussed later), with expression levels in the BHK cells as high as after time consuming amplification and the use of a bovine papillomavirus-based episomal vector (discussed later) in CHO cells.

1.2.2 Increasing gene copy number: use of viral replicons

As protein expression levels also depends on the copy number of the integrated DNA in the transfected cells, other approaches to optimise expression level has concentrated on increasing the vector copy number in stably transfected cells:

1.2.2.1 viral replicons

A viral replicon may be included in the expression vectors so that once inside the cell, it is capable of replicating episomally, provided the appropriate *trans*-acting factors are present either endogenously in the host cell genome or in the vector. The gene to be expressed is replicated to a correspondingly high copy number, leading to an increased amount of product-specific mRNA and hence to increased foreign protein synthesis. Different viral replicons work with different efficiencies, resulting in different expression vector copy number. Vectors containing the SV40 replicon replicate to very high copy number (> 10^4 copies/cell; Mellon *et al.*, 1981) in cells that express viral T antigen, for example COS cells (an african green monkey cell line, CV-1, transformed with an origin-deficient SV40 virus). However, the high viral copy number leads to cell death after 3 or 4 days, so protein production is only transient. As an example of transient recombinant antibody expression, Whittle *et al.*, (1987) expressed the chimaeric B72.3 antibody in COS cells under the transcriptional control

of the hCMV MIE promoter, with levels of antibody synthesis reaching 0.1-1 microgram/ml in the culture medium after 2-3 days. Vectors based on the vaccinia virus (VV), another high-copy replicon, have been used for the expression of many proteins, including the humanised antibody Campath-1H (Carroll *et al.*, 1992).

1.2.2.2 Selection for the integration of multiple copy number following transfection

Another approach to increase the vector copy number is to use plasmid expression vectors which contain all or part of the bovine papillomavirus (BPV) or Epstein-Barr virus (EBV) genome. The BPV or EBV fragment allows the vector to replicate to 50-100 copies per cell (and do not generally cause cell death; Sugden *et al.*, 1985), resulting in high levels of constitutive expression of the inserted genes. These vectors can be used to generate stably transfected cell lines that are continuously expressing more modest levels of protein. However, such replicative expression vectors are not employed in current commercial systems for stable high-level production of proteins in mammalian cells because the maximum levels of expression obtained are considerably lower than those using integrated amplifiable (Section 1.2.2.3) vectors, as shown for example immunoglobulins genes (Baker *et al.*, 1988) and possibly also for safety reasons. In addition, instability of expression has been reported and the viral-based vectors have limited host ranges, e.g. BPV will replicate in mouse but not in human cells.

An alternative method (Niwa *et al.*, 1991) used a bovine papillomavirus-based vector, capable of integrating into the genome, which contained a mutant neomycin resistance gene (which lowers the resistance to G418) under the transcriptional control of a weak promoter. Transfectants were selected at high concentrations of G418 (800μg/ml) so that only those clones containing high copy number of the vector (>300 vector copies integrated) could survive the selection by high concentration of G418. Using this approach, the expression levels of human interleukin-2 (IL-2), under the transcriptional control of human β-actin promoter, in mouse L cells and CHO cells were as high as previously achieved using the co-amplification strategy (see section 1.2.2.3).

A different approach described by Kaufman *et al.*, (1987) involves the use of dicistronic mRNA expression vectors in which the gene of interest is proximal to a selectable marker gene. The second open reading frame (ORF) is translated with an

approximately 100 fold lower efficiency. Cells therefore require higher levels of the dicistronic mRNA to survive selection and consequently selected clones will synthesise high levels of the desired protein encoded by the 5' ORF.

1.2.2.3 Use of an amplifiable, selectable marker gene

Another approach to achieve high levels of protein expression exploits the natural phenomenon of gene amplification that occurs in mammalian cells (reviewed in Stark and Wahl, 1984). If cells are cultured in increasing amounts of specific toxic drugs, resistant clones can be selected. This is sometimes due to a spontaneous mutation in the cell's drug transport pathway particularly at high concentration of selective agent. In some cases, however, it may be due to an overproduction of an essential enzyme whose activity is non-competitively inhibited by the drug, usually caused by an increase in copy number of the gene encoding the enzyme. Since the region of DNA amplified can be as large as 1000 kb (Stark *et al.*, 1989) a closely linked non-selectable gene, i.e. the gene of interest, is co-amplified, resulting in increased copy number and subsequently increased levels of protein expression. Amplified DNA may be either episomal as circular, acentromeric structures known as double minute chromosomes (DM), or chromosomally integrated forming a homogeneously staining region (HSR).

A widely used amplification system uses the cytotoxic drug methotrexate (MTX) to select for the amplification of an exogenously introduced dihydrofolate reductase (DHFR) gene in a CHO cell line that lacks endogenous DHFR activity (reviewed in Bebbington and Hentschel, 1987). This strategy was first demonstrated by Kaufman and Sharp (1982) who achieved high levels of SV40 T antigen expression (about 10% total cell protein). Page and Sydenham (1991) used this system for the production of a humanised antibody, Campath-1H. They co-transfected CHO-DHFR$^-$ cells with vectors containing the heavy and light chain cDNAs under the control of human β-actin regulatory sequences, with the *dhfr* gene on the light chain construct. Following two rounds of amplification, stable clones expressing up to 50 µg/ml/day of functional antibody were obtained, compared to the pre-amplification transfected pool at 0.5 µ g/ml. The system has also been used with mammalian host cells that contain an active endogenous *dhfr* gene, using a mutant DHFR selectable marker gene with increased resistance to methotrexate so that amplification of the transfected gene is favoured over that of the endogenous gene. Dorai and Moore (1987) used such a mutant *dhfr* gene when they expressed the B6.2 chimaeric antibody in a mouse myeloma cell line,

Sp2/0, with antibody chains under the control of Ig promoter/enhancer elements. They achieved a 25 fold increase in expression, from ~1 to 25 μg/10^6 cells/day, following 3 to 6 months selection with increasing MTX concentration. The length of time required for amplification makes this approach not feasible for transient expression and also the increased in expression level is very poor compared to the Page and Sydenham study. Furthermore, much higher levels of MTX were required to achieve amplification (500 μM compared to 1 μM in DHFR-deficient CHO cells used in the Page and Sydenham study).

A limitation to the DHFR/MTX system is that both DMs and HSRs amplified DNA may be unstable in the absence of selective pressure. Weidle *et al.*, (1988), studied three CHO cell lines containing chromosomally integrated co-amplified *dhfr* and *tPA* genes grown for three months in the presence and absence of MTX. In all three amplified clones, HSRs were gradually lost in the absence of selective pressure, with an accompanying decrease in tPA expression level. They proposed that this was due to homologous recombination between the tandemly arranged amplified vector copies. Similarly, Page and Sydenham (1991) found that antibody yields of an amplified clone grown in the absence of MTX remained high for 6 weeks but then gradually declined. Instability after prolonged culture in the presence of MTX has also been reported (Pallavacini *et al.*, 1990).

An alternative system has been described with the glutamine synthetase (GS) gene as the selectable, amplifiable marker gene and methionine sulphoximine (MSX) as its inhibitor. Glutamine synthetase is essential for glutamine synthesis (from glutamate and ammonia) and cells that lack the enzyme have a requirement for glutamine in the growth medium. The introduction of the GS gene into the cell provides a selectable marker allowing transfected cells to grow in the absence of glutamine. It is most often used with the GS deficient mouse myeloma cell line, NSO (Bebbington *et al.*, 1992), but may be used with cells that contain an endogenous copy of the GS gene, for example CHO cells (Cockett *et al.*, 1990), although higher levels of MSX are required for the latter. Bebbington *et al.*, (1992) used the GS amplification system for expression of the chimaeric B72.3 antibody in NSO myeloma cells. After one round of amplification a clone was isolated that secreted 9.5 μg/10^6 cells/day, compared to the unamplified productivity of 1.5 μg/10^6 cells/day.

The two amplification systems show markedly different characteristics; for example, amplification is generally necessary to achieve optimal levels of expression with the CHO/DHFR system. Gene copy number in the amplified cell lines is very high in the

CHO/DHFR system, with 100-1000 copies of the *dhfr* gene, whereas only up to 2 gene copies are found in the MSX-amplified lines. A higher degree of stability has been reported for the NSO/GS amplified lines, possibly due to the lower amplified gene copy number. Bebbington *et al.*, (1992) found that the production rate of one GS-amplified clone was constant, when cultured both in the presence and absence of MSX, for at least 30 generations. The two amplification systems were compared by Peakman *et al.*, (1994) for the production of a humanised anti-CD4 antibody. They found that the heavy chain gene and marker gene copy number was much higher in the MTX-amplified lines but that mRNA levels of the heavy chain were not correspondingly higher. As a result, the levels of expression from isolated clones from each amplification system were comparable at ~30 $\mu g/10^6$ cells/day. Similar yields, i.e. ~200µg/ml in spent culture, are routinely achieved using the NSO/GS system for the expression of a range of humanised antibodies (C. Rossmann, Wellcome Research Labs [now GWR&D], pers. comm.).

The ability to scale-up the expression system is important for the production of therapeutic proteins (e.g. antibody). It has been suggested that CHO cells are less adaptable to fermenter conditions and final yields tend to be reduced, whereas NSO cells are well suited to fermenter conditions, growing at high density in suspension (up to 10^7 cells/ml). With optimised culture conditions, using serum-free medium and nutrient concentrations, Celltech (Slough, UK) have increased cell density and productivity to achieve recombinant antibody yields of hundreds of micrograms per ml of culture medium. With such productivity, a single 2000 L fermentor can produce tens of kgs per year (necessary if the antibody is to be used therapeutically). However, a problem associated with the use of such fermenters for long term production is that post-translational modifications such as glycosylation does not keep pace with the rate of protein synthesis and thus affects the activity of the therapeutic protein.

1.2.2.4 *In vitro* amplification

Another strategy has been described to increase the copy number of the gene encoding the protein of interest *in vitro*, i.e. prior to transfection into the mammalian host cell (Ikeda *et al.*, 1988; Takeshita *et al.*, 1988). The foreign gene expression cassette is cloned into a cosmid vector at a high cassette:vector ratio, to favour the cloning of multiple copies of the expression cassette. Integral to the method is the use of a restriction enzyme that generates single-stranded cohesive termini that are not

palindromic (SfiI in the case of Ikeda *et al.*, and BstXI in the Takeshita *et al.*, study). Such a restriction site, located at either end of the fragment containing the expression cassette, ensures that the cloned multiple copies ligate together in a head-to-tail orientation only. This is important because inverted repeats, resulting from head-to-head or tail-to-tail ligation, would be unstable in a bacterial host cell (Sadler *et al.*, 1978). The cosmid clone containing the expression cassette multimer is then transfected into a mammalian host cell, where the increased gene copy number can lead to an increased expression level.

Although, the various approaches described above can improve expression of recombinant proteins, the regulation of the heterologous genes at the transcriptional level is still influenced by chromosomal "position effects". It would be advantageous to overcome this position effect and still maintain high protein expression level. The following section describes strategies that has been applied to overcome chromosomal position effects.

1.2.3 Overcoming position effects

As previously mentioned, DNA transfected into mammalian cells integrates essentially at random into the host cell genome and the positions of integration can affect the level of expression of transfected DNA. These chromosomal "position effects" have been observed by several investigators (e.g. Yoshimura and Chaffin, 1987) and in several cell types (Reff and Pfarr, 1992) and it means that many clones must be screened from a single transfection in order to isolate a high-expressing clone. However, extensive screening may be a very costly exercise (especially for the large scale production of recombinant proteins), it may be necessary to reduce the number of individual clones that have to be screened and achieve an equivalent and a more reproducible gene expression level.

1.2.3.1 Transfection with a locus control region

One approach to overcome this variability uses sequences that alleviate chromosomal position effect. For example, a locus control region (LCR): this is a genetic element that contains dominant regulatory sequences that specify position-independent expression. The position independent expression induced by LCR is thought to reflect the inherent ability of the LCR to establish an open chromatin domain and permit

regulatory factors, perhaps those mediating nuclear matrix attachment to gain access to individual genes.

In the LCR/MEL expression system described by Needham *et al.*, (1995), high-level and position-independent expression of heterologous cDNAs in murine erythroleukemia (MEL) cells was observed under the transcriptional control of the human β-globin promoter linked to the human globin LCR. A vector was constructed containing the human globin LCR upstream of the foreign cDNA expression cassette, which uses the human β-globin promoter. These authors expressed the human calcitonin receptor, a seven-transmembrane helix receptor, in this system achieving expression levels higher than previously achieved using an amplified CHO line and the baculovirus/insect cell system. Some studies have reported that the levels of expression obtained with the use of such LCR sequences are not higher than obtained following random integration and extensive screening (Grosveld *et al.*, 1987; Phi-Van *et al.*, 1990). However, the individual clones have more reproducible levels of gene expression following the use of one of these position-independent elements, and hence may decrease the number of individual clones that need to be screened to identify a high-expressing clone.

1.2.3.2 Targeted integration using Site-specific recombination system

Another approach that would eliminate the influence of chromosomal position effects on expression levels uses conservative site-specific recombination systems for the targeting of foreign DNA to a pre-selected site in mammalian genome.

Whereas integration of LCR based vectors is still random, use of site-specific recombination system will target integration of heterologous genes into a previously determined chromosomal site. Such a strategy would be advantageous if the site to which the foreign gene is to be targeted is favourable for high expression leading to the isolation of transfected cell lines expressing relatively constant, high levels of recombinant protein. This approach will create the potential for the construction of cell lines expressing high, medium or low and/or cell cycle specific induction of protein expression level and promises a general application for the large scale production of recombinant proteins.

One such conservative site-specific recombination system, the Cre-loxP, encoded by bacteriophage P1 (described in the following section) has been used extensively to

effect a wide range of precise modifications to DNA molecules both *in vitro* and *in vivo* and is exploited in this thesis as a tool that would allow the reproducible expression levels of recombinant proteins in mammalian cells.

1.3 The Bacteriophage P1 Cre-loxP site-specific recombination system

Bacteriophage P1 was recognised as one of three temperate phages harboured by lysogenic *Escherichia coli* (*E. coli*) by Bertani (1951). The three phages, namely P1, P2 and P3, were characterised by their plaque sizes on a strain of *Shigella dysenteriae*, respectively; small, large and variable. Of the three phages, P3 has received the least attention because preparations of it appeared markedly unstable. It is closely related to P2, with which it forms viable recombinants (Bertani and Six, 1988). Bertani and colleague concentrated their attention on P2 for the reason that its plaque characteristics on *E. coli* made it relatively easy to study away from *Shigella*, unlike the original isolate of P1. Phage P1 might have remained as obscure as is P3, were it not for the discovery by Lennox (1955) of P1-mediated generalised transduction between strains of *E. coli* and *Shigella* and the discovery by Lederberg (1957) of P1-mediated modification and restriction. Characterisation of P1 was instrumental in bringing about the age of genetic engineering, even if the P1 restriction nuclease has not been useful to genetic engineers. P1 as a transducing phage continues to be a key element for gross genetic manipulations in bacteria.

Ikeda and Tomizawa (1968) demonstrated that P1 prophage exists extrachromosomally as a plasmid and that unlike, for example, bacteriophage lambda (λ), P1 lysogeny does not require integration into the bacterial chromosome. As P1 prophage are not maintained by passive replication with the host chromosome (c.f. λ) but are stable and replicate at a low copy number, this suggested that P1 should be actively partitioned between daughter cells rather than passively (and randomly) distributed by diffusion.

After a considerable latent period, the findings of Ikeda and Tomizawa provided a new impetus and a new direction to P1 biology. Genetic studies of P1 was initiated by Scott (1968) and extended by Walker and Walker (1975, 1976b), and Bachi and Amber (1977). DNA isolated from phage particles is double-stranded, linear with about 12% terminal redundancy, and cyclically permuted. Its molecular weight is about 6×10^4 kDa (Ikeda and Tomizawa, 1968). In contrast to other phages with this

type of DNA organisation (P22, T4), whose genetic maps are circular, the genetic map of P1 is surprisingly linear despite the phage DNA being cyclically permuted. These findings suggested that either a large segment of DNA at the ends of the P1 genetic map is devoid of available genetic markers, or this region of the P1 genome contains a hot spot for genetic recombination. The first possibility was eliminated by mapping studies that located markers present at both ends of the P1 genetic map on a unique 6.5 kb *Eco*RI restriction fragment (*Eco*RI-7; Sternberg, 1978; 1979). The second possibility was investigated using this *Eco*RI-7 fragment (which most likely contained at least some elements of this recombination system) in genetic crosses to promote recombination between flanking λ markers under conditions in which both λ site-specific recombination system and host recombination systems were inactivated by mutation (Sternberg and Hamilton, 1981a). Sternberg and Hamilton (1981a) studies showed that the *Eco*RI-7 fragment contained all the essential components of a recombination system. Two components were identified, a site (loxP) that must be present in both partners of a genetic cross for efficient recombination, and a recombinase gene (*cre*), whose protein product, Cre can act in trans to promote recombination.

1.3.1 The role of Cre-loxP recombination system in the life cycle of bacteriophage P1

Various essential and diverse roles has been attributed to the Cre-loxP recombination system in the life cycle of P1.

1.3.1.1 Role of Cre-loxP recombination in the cyclisation of P1 DNA

The Cre-loxP recombination system is essential for the cyclisation of newly injected terminally redundant virion DNA. Segev *et al.*, (1980), showed that circular P1 molecules could be detected as early as five minutes after injection of linearised P1 DNA into *E. coli*, well before the DNA is replicated. The importance of the cyclisation event is accentuated by the almost complete failure of P1 to form lytically or establish lysogeny in the absence of the Cre-loxP recombination system (Sternberg *et al.*, 1986). Presumably, if the injected P1 DNA is not cyclised, it is rapidly degraded from its ends by cellular nucleases such as the products of the host *rec*BCD genes. The properties of P1 *cre* mutants lend support to the proposed role of the Cre-loxP system in the cyclisation reaction. These mutants lysogenise *rec*A bacteria 10-25

times less efficiently than does a *cre*$^+$ phage, but they are indistinguishable from *cre*$^+$ phage in their ability to lysogenize *rec*A$^+$ bacteria (Schultz *et al.*, 1983). Because *cre* lysogens are no less immune than *cre*$^+$ lysogens (Sternberg, unpublished) the reduced lysogenisation frequency in the *rec*A cells probably reflects the failure of the *cre* DNA to cyclise. Its ability to cyclise in *rec*A$^+$ bacteria indicates that other systems dependent on RecA function can replace Cre-loxP for cyclisation (Austin *et al.*, 1981). The most obvious alternative to the Cre-loxP system for cyclising P1 DNA is the general recombination system of the host. Theoretically, it could generate P1 circles by recombination at any position within the terminally redundant region of an injected P1 DNA molecule. An indication that this alternative might not function comes from an experiment in which chloramphenicol-treated cells were shown not to be able to cyclise injected P1 DNA despite the fact that they should still be recombination proficient (Segev and Cohen, 1981). It is possible that P1 encodes a protein that enhances bacterial generalised recombination and permits injected P1 DNA to cyclise. However, a P1 cyclisation enhancement gene has not yet been found.

1.3.1.2 Role of Cre-loxP in P1 vegetative growth and lysogeny in host cell

P1 DNA that has been injected into a sensitive bacterium, that has circularised by recombination, can induce one of two states, namely lysis or lysogeny. If conditions such as temperature are favourable, prophage establishment occurs. However, the prophage state is only as stable as the conditions that favours it. Sternberg *et al.*, (1986) showed that Cre is not required for viral vegetative growth and lysogeny in a *rec*A$^+$ host but is required for both processes in a *rec*A$^-$ host. The requirement for vegetative growth suggests that Cre also has a role to play in the viral lytic cycle after the viral DNA has been cyclised.

1.3.1.3 Role of Cre-loxP in maintenance of unit-copy P1 plasmid prophage

The Cre-loxP recombination system facilitates the proper partition of P1 plasmids in host cells. The content of P1-specific DNA in lysogenic *E. coli* is estimated to be 1.0 per bacterial chromosome (Ikeda and Tomizawa, 1968); for this reason P1 prophage is referred to as a unit copy plasmid or a plasmid under stringent control. Both host and plasmid functions contribute to copy number control as evident from the existence of *E. coli* mutants (Cress and Kline, 1976) and P1 mutants (Sternberg 1978). The role of Cre-loxP recombination system is to segregate P1 dimers (which are formed by

homologous recombination between two daughter molecules after replication) into monomers (Figure 1.1; Austin *et al.*, 1981). Because the daughter molecules are physically linked in the dimer, they cannot be segregated to separate cells at cell division. Recombination between the two loxP sites present in the dimer efficiently regenerates the daughter monomer circles so that they can be faithfully segregated (Sternberg *et al.*, 1980; Austin *et al.*, 1981). The Cre-loxP system plays varied, and normally minor roles in the life cycle of P1, but owing to its intrinsic interest has been the object of intensive study.

1.3.2 Cre is a member of the int family of recombinases

A combination of methods for detecting distant relationships in protein primary sequences has been used to show that Cre recombinase belongs to the Int (for integrase) family of site-specific recombinases which includes the site-specific recombination proteins encoded by bacteriophage λ, ϕ80, P22, P2, 186 and P4 (Argos *et al.*, 1986). All seven of the recombinase proteins are basic and approximately the same size (336-440 amino acids), but there is very little amino acid sequence homology within this group. Despite this diversity, all of the recombinases can be aligned in their C-terminal halves. A 40-residue region near the C-terminus is particularly well conserved in all the proteins and is homologous to a region near the C-terminus of the yeast 2μ plasmid Flp protein (McLeod *et al.*, 1984). This family of recombinases does not appear to be related to any other site-specific recombinases. Three positions are completely conserved within this family: histidine, arginine and tyrosine are found at respective alignment positions 396, 399 and 433 within the well conserved C-terminal region. It is speculated that these residues contribute to the active site of these proteins. It is possible that the tyrosine (Y-433) forms a transient covalent linkage to DNA during strand cleavage and rejoining.

Although there are important differences among these recombination systems such as their requirement for accessory proteins in recombination, the genetic and biochemical data suggest that the basic features of recombination, particularly the strand exchange event are similar. The Int family of recombinases are characterised by a strand exchange mechanism that requires no DNA synthesis or high energy co-factor(s). As in the case of several DNA topoisomerases, the phosphodiester bond energy is conserved in a phospho-protein linkage during strand cleavage and religation (Craig and Nash, 1983; Andrews *et al.*, 1985). The site-specific recombination systems of bacteriophages ϕ80 and P22 appear closely related to that of λ in their requirement

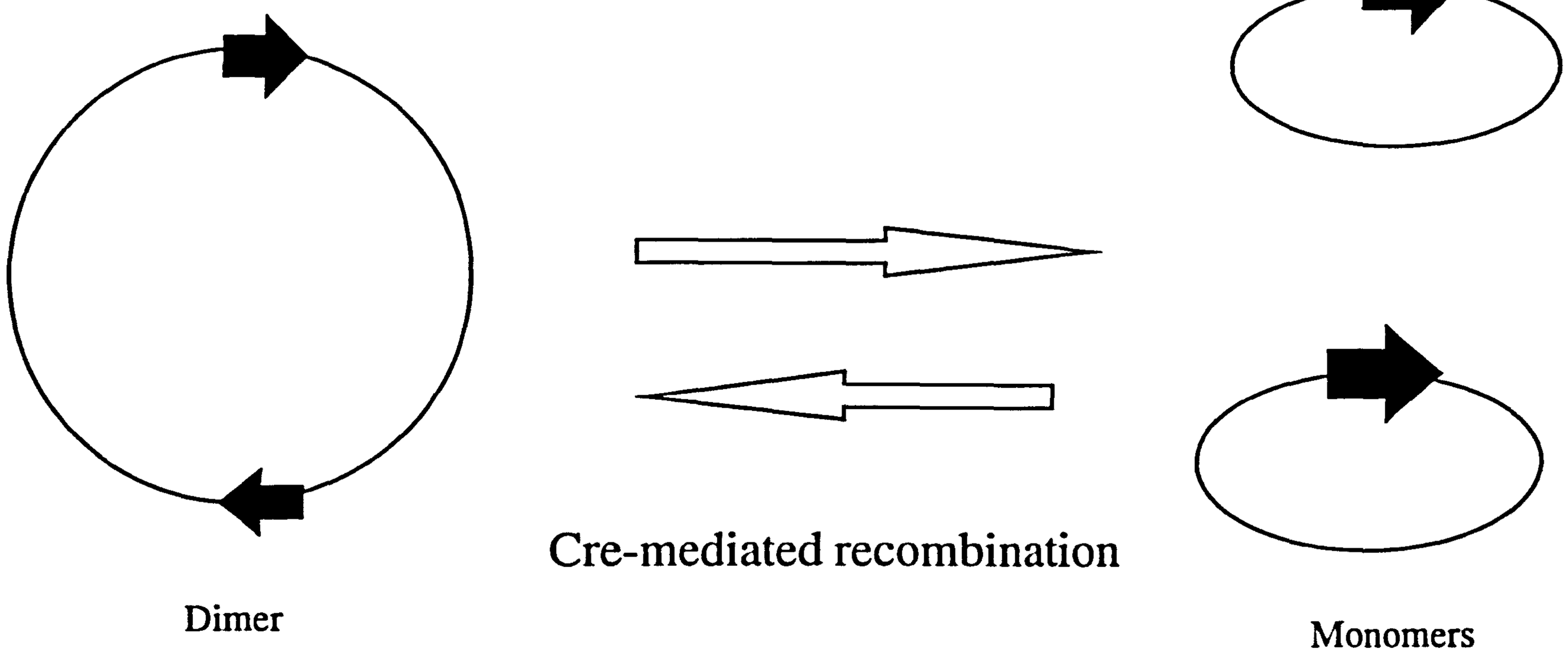


Figure 1.1- Schematic representation of Cre mediated recombination between two loxP sites (black arrows). Cre recombinase can mediate intramolecular recombination between two loxP sites (in direct repeat) on a dimeric molecule, resolving them into two monomers. Conversely Cre recombinase can mediate intermolecular recombination between loxP sites on two monomers converting them back to the original substrate.

for integration host factor protein (IHF) and for an excisionase protein (Xis) during excisive recombination (Leong *et al.*, 1985). P2 requires a protein (Cox) in addition to Int for prophage excision (Lindahl *et al.*, 1972) suggesting that the directional controls are relevant in P2 as well. It is interesting that Cre and Flp, which are not under directional control, are also the two most distant from the other members of the Int family. Cre and Flp are particularly closely related to each other (functionally), as demonstrated by the fact that Cre binds to the Flp recombination site (albeit with a 20-fold lower affinity; Abremski, K. and Jayaram, M. unpublished observation). In spite of their apparent close relationship, Flp and Cre align only over the 40-residue segment near the C-terminus.

1.3.3 The Cre recombinase

Cre, a 38.5 kDa protein is encoded by 343 amino acids and separated from the loxP site on the P1 genome by a 434 base-pair (bp) region that contains a 73 amino acid open reading frame, ORF1 (Sternberg *et al.*, 1981). The *cre* gene and ORF1 are oriented with their amino-terminal ends proximal to loxP. Three promoters have been identified as located upstream of the *cre* structural gene. Their activities range from 7 to 10% of the activity of a galactose operon promoter and are active in P1 lysogens. The only evidence to date for the regulation of *cre* expression is that the promoter furthest from *cre*, pR1, contains two *dam* methylation sites (5' -G-A-T-C- 3') in its -35 region (with respect to its translational initiation site) which is sensitive to methylation (Hattman *et al.*, 1978). Its transcription is three to fourfold higher in a dam^- host than it is in a dam^+ host. The promoter closest to *cre*, pR3, directs the production of an RNA transcript that functions inefficiently for Cre protein synthesis because it lacks a ribosome recognition site (Sternberg and Hamilton, 1981a; Abremski *et al.*, 1983). The transcriptional activity of the three Cre promoters is not affected by other proteins expressed by the P1 prophage, including the cI repressor protein.

In solution, the Cre protein exists as a monomer. A limited chymotryptic digest of Cre results in two fragments of sizes 25 and 13.5 kDa. The sequence of the amino terminus of the purified 25 kDa peptide demonstrates that this peptide represents the carboxyl-terminal portion of the Cre protein. The 25 kDa peptide is capable of specific binding to the loxP site but binds with a lower affinity than does the wild type Cre protein.

1.3.4 The loxP site

A functional loxP site is quite small, comprising a 34 bp sequence (Abremski *et al.*, 1983). DNA footprinting analysis with purified Cre (Hoess and Abremski, 1984) as well as deletion analysis (Hoess *et al.*, 1982) has revealed that the loxP site consists of two 13 bp inverted repeats separated by an 8 bp spacer region (Figure 1.2). Abremski and Hoess (1984) showed that each inverted repeat and contiguous 4 bp of the spacer region comprises a binding domain for Cre. The absolute stoichiometry of the Cre-loxP complex has been determined to be one molecule of Cre bound per inverted repeat, or two molecules per loxP site (Mack *et al.*, 1992).

From the DNA footprinting and deletion analysis, it can be determined that the 8 bp spacer region is the only asymmetric feature of the loxP site. This has important consequences for the outcome of recombination between loxP sites. When two loxP sites on the same DNA molecule are directly repeated, recombination results in the excision of the DNA between the sites. If, however, the sites are in an inverted orientation with respect to each other then, when recombination occurs, the DNA between the sites undergoes inversion rather than excision. Of particular interest to a number of groups is how the asymmetry of the 8 bp spacer region confer directionality on the site? Is the directionality achieved by DNA-DNA homology between the two recombining sites?

To address these questions a number of mutant loxP sites containing alterations in the spacer region were made (Hoess *et al.*, 1984). The activity of these mutants during recombination led to a number of important observations regarding the mechanism of recombination. First, sequence homology in the 6 bp overlap region is a requirement for recombination between loxP sites (a mutant site will not recombine with a wild-type site, but a mutant site will recombine with itself). These observations are analogous to those found for the overlap regions of the λ att site (Weisberg *et al.*, 1983), the yeast 2μ Flp site (Senecoff *et al.*, 1988), and the *Salmonella* Hin system (Johnson and Simon, 1985). Thus, it seems that the requirement for sequence homology appears common to many site-specific recombination systems.

A second observation is that homology is not sufficient for recombination if the mutation is at position +1 (position number of bases with respect to axis of dyad symmetry in the spacer region). The +1 position then appears critical for recombination suggesting that perhaps it is a base where there is an essential specific

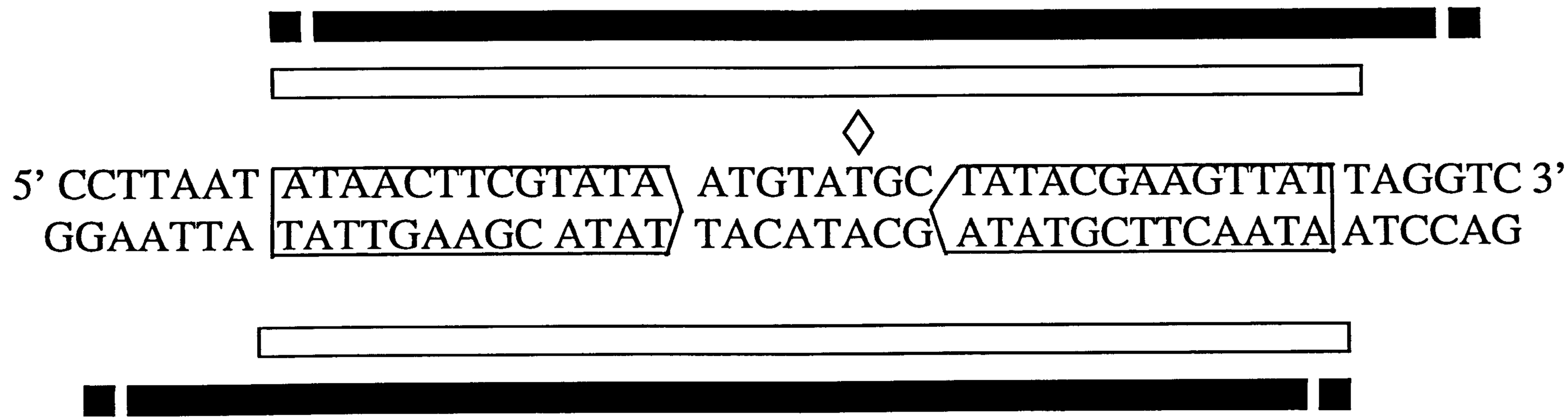


Figure 1.2- Sequence of loxP protected by Cre recombinase in bacteriophage P1 in a DNA-protein footprinting experiments.The sequence of loxP is shown with the patterns of protection by Cre against DNase 1 (▮) or NCS (▯) cleavage indicated in each strand. The small squares (■) are at positions where DNase does not cut even in the absence of Cre, so that protection of those bases are uncertain. The diamond tool (◇) indicates the cleavage by NCS at the T position in the upper strand. The axis of dyad symmetry is considered to be at the midpoint of the spacer region between 5' T-A 3' of the top strand.The 13 bp inverted repeats are shown in boxes with arrowheads flanking the spacer region.

contact between nucleic acid and protein. Footprinting data suggest that at least some of the spacer region must be in contact with Cre (Hoess *et al.*, 1984). However, this was done using DNaseI and not hydroxyl radicals (which is more precise for footprinting). A second possibility is that certain sequences themselves play a crucial physical role during the recombination process. Drew *et al.*, (1985) proposed that the TpA sequence in the spacer region is predisposed to melting. This offers the possibility that the TpA sequence is critical for unwinding of the loxP site either during the pairing or the strand exchange step of recombination. Additional evidence shows that the sequence of the spacer region dictates the directionality of loxP recombination. By creating a loxP site with a completely symmetric spacer region, recombining sites now are capable of both excision and inversion in the same construct both occurring with approximately equal efficiency (Hoess *et al.*, 1984).

At a very low frequency, P1 can integrate its DNA at a specific site called loxB, in the E. coli chromosome (Sternberg *et al.*, 1981), creating two hybrid sites, termed loxL and loxR which flank the integrated prophage. These four sites, loxP, loxB, loxL and loxR, have been isolated on λ transducing phages and their DNA sequences has been determined (Hoess *et al.*, 1982). The crossover point in loxP X loxB recombination occurs in the 8 bp spacer region, and deletion mutations that remove either of the repeats plus a portion of the spacer inactivates the site. The bacterial loxB site could be important for the transient integration of P1 DNA in the production of generalised transducing phage.

1.3.5 Cryptic loxP site(s)

Because the loxP site of phage P1 is 34 bp in size, the natural occurrence of this exact sequence is unlikely in any eukaryotic genome. However, related sequences may exist in eukaryotic genomes that could recombine at low efficiency with an authentic loxP site. Such cryptic loxP sites have been identified in the yeast genome (Sauer, 1992). Because of the greater complexity of the human and mouse genomes compared with yeast, an analogues situation is likely to exist in these organisms. This implies the possibility of Cre-mediated illegitimate recombination in mammalian cells.

1.3.6 The mechanism of Cre-loxP recombination reaction

Hoess *et al.*, (1985) showed that the full inverted repeats and the spacer region are required for recombination to occur. One binding domain is insufficient. When DNA

containing the loxP site was incubated with Cre, specific cleavages occurred within the spacer region, creating a six base-pair staggered cut. The cuts were centred on the axis of dyad symmetry of the loxP site, resulting in a 5' protruding terminus.

5'A^ T G T A T G C3'
3' T A C A T A C^G5'

At the point of cleavage, Cre becomes covalently attached to a 3' phosphate group, and produces a free 5' hydroxyl group (Craig and Nash, 1983). A series of experiments done by Hoess *et al.*, (1985) in which a radioactively labelled loxP site was recombined with an unlabelled loxP site located the point at which strand exchange takes place during recombination. The point of strand exchange coincided with the sites at which Cre cleavage of the DNA backbone had been detected. Cuts in the recombining site of each strand were made opposite one another and strand exchange followed cleavage. The cleaved DNA must then pair with the complementary single-stranded sequences of the other loxP partner in the recombination reaction. The exchange of these strands results in a Holliday structure (Holliday, 1964), which, by virtue of the heterology surrounding it, would be immobile. The recombinase resolves such structures by cleaving the Holliday junction into the resulting heteroduplex products.

The breakage and rejoining of partners during this recombination process must be rapid because it has not been possible to detect intermediates of the reaction at least when using wild-type system. Several mutant Cre proteins have been isolated that produce significant amounts of a possible intermediate product of the recombination reaction. A large body of evidence (both genetic and physical) has implicated the Holliday structure as an intermediate in recombination (Hoess *et al.*, 1987). Perhaps most convincing is the direct observation of such intermediates by electron microscopy.

The fundamental chemistry of recombination is phosphoryl transfer involving the translocation of a phosphodiester within a nucleic acid chain or between two nucleic acid chains (Jayaram, 1994). A nucleophile derived from the Cre recombinase enzyme breaks the phosphodiester bond, forming a covalent link between the DNA and the enzyme. The cleavage reveals a sugar hydroxyl that provides the active nucleophile for the second transesterification step (strand exchange). The strand cleavage is catalysed by the active site tyrosine (Y-433); the linkage between protein and DNA is through the 3'-phosphate and the strand exchange is initiated by the 5'-hydroxyl of the

nicked partner DNA. The whole recombination reaction is initiated by one round of reciprocal single-strand exchanges, branch migration of the resulting Holliday junction and a second round of single-strand exchanges complete the recombination reaction (this general mechanism is followed by recombinases belonging to the Int and Resolvase/Invertase families) Craig *et al.*, (1988).

1.3.7 Factors required for the Cre-loxP reaction

Known host recombination factors are not necessary for Cre-loxP recombination. Recombination measured in bacterial strains containing mutations that affect either homologous recombination (*rec*AB) or λ site-specific recombination (*him*A, *him*B; Miller *et al.*, 1978) do not significantly affect the efficiency of the P1 site-specific recombination (Sternberg and Hamilton, 1981b).

The *in-vitro* P1 recombination system requires the addition of only a Cre-containing extract and divalent or polyvalent cations (e.g. Mg^{2+} or spermidine) without any external energy cofactor(s) (e.g. ATP), suggesting that no external energy source is needed (Sternberg and Hamilton, 1981b; Abremski *et al.*, 1983). Spermidine may act to condense the DNA, while the effect of Mg^{2+} on the reaction might affect the protein's ionic interaction with DNA.

1.3.8 Topology of DNA substrate required for the recombination reaction

A number of λ and plasmid substrates containing two loxP sites have been constructed as model systems. Using these substrates it has been shown, both *in vivo* and *in vitro* that a fully functional loxP site is composed of no more than 60 bp (Abremski *et al.*, 1983; Hoess and Abremski 1985; Sternberg and Hamilton 1981b), suggesting that all the informational content needed for recombination must reside within these sequences.

The conformation of the DNA substrate containing the loxP site does not appear to be important at least for the *in vitro* reaction. Thus, when an extract containing Cre is used, intramolecular recombination between loxP sites on supercoiled, nicked-circle or linear DNA occurs efficiently. In addition, the intramolecular loxP X loxP reaction occurs efficiently regardless of the orientation of the two sites relative to one another. Intermolecular recombination can occur between two supercoiled DNA molecules or

between a supercoiled/circular/linear and one supercoiled molecule or between two linear DNA molecules. When both substrates are supercoiled the resulting product is a supercoiled dimer. When one substrate is supercoiled and the other is linear, the product is a linear dimer. Intermolecular reaction appears to be less efficient than intramolecular. The intermolecular reaction occurs very rapidly; the majority of the products are formed within 2-4 minutes at 37°C. However, the extent of this reaction is never greater than 10%-15%, even after 60 minutes of incubation. In contrast, the intramolecular reaction reaches levels of 60% to 70% within a similar time course (Abremski *et al.*, 1983).

The most surprising result from the *in vitro* studies is that 50% of the products of recombination (intramolecular) between loxP sites on a supercoiled DNA substrate are present as free supercoiled circles. The ability to produce free products starting with a supercoiled substrate suggests a unique property of Cre-mediated loxP recombination (Abremski *et al.*, 1983). Either Cre or a protein in a crude extract containing Cre acts as a topoisomerase (Krasnow and Cazzarelli, 1982), so that while the initial recombinant products are all catenated the topoisomerase activity acts to decatenate these products or the decatenation could be an intrinsic part of the recombination mechanism. For instance, Cre might recognise and bind to one or both loxP sites on the DNA molecule and track along the DNA by some non-specific "electrostatic" interaction proposed by (Berg *et al.*, 1981), until another loxP site is found and recombination can take place. The net effect would be that during the tracking process the DNA strands between the loxP sites would be separated from one another effectively reducing the probability that the product molecules will be interwound. Cre could bring the loxP sites together by some mechanism that ensures the frequent separation of the two halves of the molecules. From both *in vitro* and *in vivo* experiments, loxP sites separated by a few bases work just as efficiently as those that are surrounded by nearly a kilobase of P1 DNA (Adams *et al.*, 1992).

In contrast with the *in vitro* studies, where Cre does not show strong preference for substrate topology, site connectivity or pathway for synapsis, *in vivo* Cre is a much more specific enzyme (Adams *et al.*, 1992). *In vivo*, Cre generates simple products and is stimulated by negative supercoiling. The causes of these differences in Cre behaviour *in vivo* is unknown. DNA structure and ionic conditions *in vivo* or the influence of a host protein factor are possible candidates.

The Cre-loxP recombination event has been studied in great detail by using an *in vitro* system that efficiently shows recombination between two loxP sites (Abremski *et al.*,

1983; Hoess and Abremski, 1985; Sternberg and Hamilton, 1981b). At equilibrium, Cre efficiently converts 70% of DNA substrate to products and appears to act stoichiometrically. The purified Cre protein binds to loxP containing DNA and makes complexes that are resistant to heparin (heparin competes for any non-specific interaction between protein and DNA; Abremski and Hoess, [1984]). This result is a further prove that Cre makes specific complexes with the loxP site.

This characteristics of the Cre-loxP site-specific recombination system suggests a new strategy for introducing DNA into the genome of eukaryotic cells in addition to the already well characterised phage λ site-specific recombination system. In this system recombinant λ vectors carrying foreign DNA can integrate into the genome of *E. coli* at a specific location directed by the λ int protein. The advantages of this phage λ site-specific recombination have long been appreciated by bacterial geneticists (e.g., phage λ replacement vectors that can contain up to 23 Kb of contiguous DNA have been used for the isolation and analysis of genomic DNA sequences, Sambrook *et al.*, 1989; Meese *et al.*, 1990). Site-specific recombination in higher eukaryotic cells would be of immense importance for the precise DNA manipulations in the genome. Functional heterologous expression of site-specific DNA recombinases in eukaryotic cells has been shown not only for the Cre protein of bacteriophage (Sauer and Henderson, 1988), but also the Flp protein of *Saccharomyces cerevisiae* (Golic and Lindquist, 1988; O'Gorman *et al.*, 1991) and the pSRI recombinase of *Zygosaccharomyces rouxii* (Matsuzaki *et al.*, 1990).

1.4 Applications of the Cre-loxP site-specific recombination system

That the Cre-loxP site-specific recombination system of bacteriophage P1 can function in an efficient manner in eukaryotes was initially shown in the yeast *Saccharomyces cerevisiae* (Sauer, 1987). The yeast selectable marker, *leu*2 gene flanked by two LoxP sites was integrated into the yeast chromosome by homologous recombination. The *cre* gene was expressed in *S. cerevisiae* by fusing the *cre* structural gene to an inducible GAL1 promoter. Induction by galactose led to the production of active Cre protein which was able to catalyse efficient and precise recombination between the two loxP sites flanking the *leu*2 gene. Results of recombination led to the excision of the *leu*2 gene (as measured by LEU^- phenotype and the resulting leucine auxotrophs). This result demonstrates that Cre protein can enter the nucleus of the yeast *S. cerevisiae* to efficiently recognise and catalyse

recombination at loxP sites introduced to native chromosomes. The precise and yet simple mechanism with which Cre mediates recombination at the loxP sites suggested a potential method for genetic manipulation of the eukaryotic genome.

1.4.1 Cre-loxP recombination in plants

Recent advances in the genetic engineering of plants have led to the production of a large variety of crop plants with agriculturally valuable traits (Goodman *et al.*, 1987; Gasser and Fraley, 1989). Notable examples are engineered plants that show enhanced tolerance to plant pathogens (Vaeck *et al.*, 1987; Johal and Briggs, 1992). However despite the scientific advances made in crop improvement, the commercialisation of genetically engineered plants has been slowed by public concerns on the issue of the environmental safety of genetically engineered organisms. One such concern is with the genes that code for antibiotic resistance. Co-transformed into the host genome as selectable markers, they serve a necessary role in the detection of genetic transformants. Hence, agriculturally desirable characteristics are invariably co-introduced with resistance to antibiotics or herbicides. The products of these selection systems are not necessarily harmful, but their presence in transgenic plants does increase the environmental uncertainty of their distribution. This uncertainty factor can have a highly negative effect on the progress of plant improvement. Gene transfer without the incorporation of antibiotic-resistance markers in the host genome should ease public concerns over the field release of transgenic organisms expressing such traits.

Dale and Ow (1991), described a strategy for engineering plants free of selectable markers. A luciferase gene was introduced into the tobacco genome using the hygromycin phosphotransferase gene (*hpt*) as a linked selectable marker which is flanked by two loxP sites. Cre recombinase was shown to excise the *hpt* in the plant genome by site-specific recombination at the loxP sites, resulting in plants which had incorporated only the required transgene. An advantage in removing the marker gene used to select transformants is that those cells can be re-transformed with additional traits in a stepwise process using the same resistance gene. This can reduce the restriction imposed on experimental designs by the limited availability of selectable markers. The concern about the environmental impact and public perception of widespread release of organisms expressing genes that confer resistance to antibiotics or herbicides can be alleviated by employing the strategy described above to remove selectable markers from the host genome. However, the Cre-loxP system leaves a

loxP site in the plant genome which may not have any effect on gene expression as long as it is not in a region crucial for transcription.

The use of Cre-loxP recombination system in conjunction with transposable elements would create the possibility to induce chromosomal deletions (von Haaren and Ow, 1993). This system, based on the mobility of one of the recombination sites introduced into the plant genome, provides the possibility to induce several forms of genome rearrangements; for instance an *Agrobacterium tumefacien*-transferred (T)-region containing a loxP site and a Ds (the maize dissociation transposable; Bancroft *et al.*, 1993) element carrying a second loxP site has been integrated into a plant genome. Transactivation of the Ds element with the Ac (the maize activation transposable element, Chuck *et al.*, 1993) transposase function results in relocation of the Ds element and the consequent separation of the two sites over different unknown distance on the same chromosome or on different chromosomes. Subsequent Cre-mediated recombination between these separated recombination sites will lead to deletion, inversion or translocation of DNA segments, dependent on the relative orientation and chromosomal location of the loxP sites.

Thus "transposition-deletion" system could allow the screening of large segments of the genome for interesting genes and may also permit the cloning of the DNA corresponding to the deleted material by the same site-specific recombination reaction *in vitro*. This methodology may provide a unique means to construct libraries of large DNA clones derived from defined parts of the genome, the phenotypic contribution of which may be displayed by the mutant carrying the deletion.

1.4.2 Cre-loxP mediated gene targeting in mammalian cells

Stable transformants of mammalian cells from standard gene transfer often show variability in expression of the introduced transgene. This occurs both from the highly variable number of copies integrated into the genome and from position effects on gene expression due to random integration (Lacy *et al.*, 1983). Hence experiments such as mutational analysis of transcriptional regulatory sequences using stable cell transformants is very difficult. Targeted integration of DNA into the genome would provide a simple solution to one of the major problems with gene transfer into mammalian cells.

Fukushige and Sauer (1992) used the Cre-loxP system to eliminate both the constraints of position effect and copy number variation on gene expression. A positive selection vector system was designed to directly select Cre-mediated DNA integration at a loxP target site previously placed into the genome of cultured mammalian cells. Initially an inactive lox-*neo*^{r} fusion gene was used as a loxP chromosomal target. Cre-mediated integration of a promoter-transcriptional initiating codon (ATG)-lox targeting construct into the chromosomal target reconstructs a functional ATG-lox-*neo*^{r} fusion gene and activates expression of a defective neomycin phosphotransferase gene which confers resistance to the neomycin analogue G418 in transfected cells which had undergo Cre-mediated targeted integration. With CHO cell lines containing this target, almost all of the selected transformants were simple single copy integrants of the target DNA. A *lac*Z reporter construct used to monitor gene expression showed that independent G418-resistant colonies from site-specific integration of the reporter gene all showed nearly identical levels of β-galactosidase activity when the reporter construct integrated at a particular chromosomal position. The same construct integrated at a second chromosomal position exhibited a slightly different level of activity, characteristic of the second position. The result showed that Cre-mediated site-specific integration can facilitate the construction of isogenic cell lines and thereby permit reproducible gene expression in stably transformed cell lines. This system could be useful for the large scale production of recombinant proteins especially if the loxP site to which the foreign gene to be targeted is favourable for high level expression. Therefore the work described in this thesis, uses a secreted alkaline phosphatase (SAP) reporter gene fused to a loxP site as the chromosomal loxP target. Thus, the expression level of SAP from independent cell lines each containing a single copy of the loxP insert should give an indication of which cell lines contain a loxP site integrated into a chromosomal region favourable for gene expression. Subsequent Cre-mediated targeted integration of exogenous DNA to a loxP site in these cell lines should give the reproducible high level expression of recombinant protein.

1.4.3 Manipulation of transgenes by Cre-loxP recombination system

Transgenic mice are now routinely generated either by direct pronuclear injection of exogenous DNA into fertilised zygotes (Palmiter and Brinster, 1986) or by injection of genetically engineered embryonic stem (ES) cells into the mouse blastocyst (Schwartzberg *et al.*, 1989).

Direct pronuclear injection results in random integration of the injected DNA into the genome and relies on the dominant nature of the transgene to give a useful phenotype (e.g. the expression of a foreign protein or of a mutated version of an endogenous gene). Manipulation of endogenous genes (gene targeting or knockout) has primarily been performed with ES cells because they can be manipulated and grown *in vitro* as a permanent cell line and then re-introduced into the mouse blastocyst to contribute to the mouse germ line. A desired goal of transgene technology is efficient and accurate manipulation of DNA sequences after their integration in the germ line. DNA recombinases that mediate integration or excision of sequences at specific recognition sites in both prokaryotic (Stark *et al.*, 1989; Sandmeier *et al.*, 1990) and eukaryotic (Broach and Hicks, 1980; Oeltinger *et al.*, 1990) systems are well suited for this purpose.

Lakso *et al.*, (1992) were the first to use the Cre-loxP system to investigate targeted activation of large tumour antigen (T. Ag.) transgenic mice. A transgenic mouse was created which contained an inactivated T. Ag. created by insertion of a specially designed stop sequence that prevents gene expression flanked by loxP sequences. By crossing the dormant T. Ag. transgenic mouse lines with Cre expressing transgenic lines, Cre-mediated recombination at the loxP sites led to the excision of the stop sequence and ensuing activation of T.Ag. The event of the site-specific recombination led to double transgenic offsprings with developed lens tumours. Since then other groups have used the Cre-loxP systems for the various manipulations of mammalian genome (described below).

As mentioned above, studies of gene function in mice often involve the analysis of embryonic stem (ES) cell-derived gene targeted mice. Such mice carry a pre-designed mutation in their germ line and are devoid of a particular gene product throughout ontogeny (Koller and Smithies, 1992; Melton, 1994). A strategy for conditional, cell-type specific gene targeting was developed by Gu *et al.*, (1994) using a combination of the conventional transgenic technique and the Cre-loxP system. Two transgenic mice were constructed, one strain in which a Cre transgene is expressed in a cell type-specific or developmentally stage-specific manner. The second strain carries the target (polβ) gene flanked by two loxP sites, which has been shown to be one of several enzymes involved in DNA repair machinery in T cells (Fry, 1983). By crossing these two strains, these authors were able to show that in T cells, ~40% of offsprings derived from an intercross were homozygous for the polβ mutation. This was the first report of a normal mouse gene being selectively knocked out in just one cell type.

Thus in principle, Cre-loxP-mediated gene targeting should allow the inactivation of any gene in any tissue at any stage of development.

The approach described by Gu *et al.*, (1994) however has a major drawback in that the function of the gene product must be deduced from the phenotype of animals that are, either in all cells or in certain cell types from an early time point on, constantly deficient throughout ontogeny for the product of the disrupted gene. A mutant organism may compensate for the loss of a gene product so that no obvious deviation from the wild-type is seen, or the organism may react to the mutation to give a complex, secondary phenotype. Moreover, if the complete loss of a gene product results in embryonic lethality (Copp, 1995), gene function at later stages of development cannot be analysed.

To overcome these limitations, Kühn *et al.*, (1995) and Zhang *et al.*, (1996) have developed methods of gene targeting that allows the inducible inactivation of a target gene in mice. The method uses an inducible promoter to control the expression of the Cre recombinase which, when activated, can catalyse site-specific recombination at loxP sites flanking a target gene in transgenic mice. This system should prove useful in mutating or deleting genes at specific times during development or in adult animals (e.g. at a given time during T-cell ontogeny or neural development).

In most gene targeting experiments only a small percentage of cells incorporate foreign DNA after transfection. Selection for a drug resistance marker contained within the targeting vector is applied in order to obtain cells into which the construct has integrated. Southern blot analysis and/or PCR are subsequently used to identify the desired homologous recombination from this initial pool of drug-resistance clones.

The most widely used forms of gene targeting (simple replacement or insertion events, Capecchi, [1989]) commonly result in the permanent introduction of positive selectable markers into the mammalian genome. There are instances, however, when the irreversible introduction of selectable markers as a result of gene targeting may be undesirable. First, the relatively small repertoire of positive selectable markers available for use in mammalian cells limits the number of mutations that can be generated in a single cell (Eglitis, 1991). Second, the disruption of both alleles of an autosomal gene usually requires the construction of two different targeting constructs bearing different selectable markers. Finally, the exogenous promoter and enhancer elements required for the expression of these selectable markers have the potential to

interfere with endogenous regulatory elements present in the vicinity of the targeted mutations.

Therefore it may sometimes be advantageous to remove the selectable markers after gene targeting. Although existing methodologies such as hit and run (Hasty *et al.*, 1991) and double replacement (Stacey *et al.*, 1994) allow for the generation of marker-free targeted alleles, the removal of selectable markers by these techniques relies on a second relatively inefficient round of homologous recombination.

Abuin and Bradley (1996) used a combination of homologous and Cre-loxP-mediated recombination to generate mouse ES cell lines carrying up to four targeted mutations and devoid of exogenous selectable markers. A cassette that contains both positive and negative selectable markers flanked by loxP sites, was used in homologous recombination and positive selection to disrupt a Rep-3 locus, a gene homologous to members of the mutS family of DNA mismatch repair genes. Cre mediated recombination and negative selection were then used to recover clones in which the cassette had been excised. The remaining allele of Rep-3 was then subjected to a second round of targeting and excision with the same construct to generate homologous, marker free cell lines. Subsequently, both alleles of mMSh2, another mutS homologue, were disrupted in the same fashion to obtain cell lines homozygous for targeted mutations at both the Rep-3 and mMSh2 loci and devoid of selectable markers.

In principle, the Cre-loxP mediated excision or recycling of selectable markers (initially shown in yeast; Sauer, 1994) after gene targeting should allow for the generation of an unlimited number of targeted mutations in a single mammalian cell line. This system may prove to be particularly useful for the study of mammalian genes whose function can be assayed at the cellular level such as DNA repair genes.

Chromosomal rearrangements are major causes of inherited human diseases and foetal loss (Epstein, 1986). Translocation (Rabitts, 1994) and loss of heterozygosity (Lakso *et al.*, 1991) are important genetic changes causally involved in neoplasia. The ability to manipulate mammalian chromosomes *in vivo* to produce translocations, inversions, and deletions with predetermined breakpoints would be a considerable technological advance. However the technology for making large-scale precisely defined alterations to the mammalian genome has not yet been established. In the mouse very few deletions are available, and those that are, were generated randomly using mutagenesis approaches such as ionising irradiation (Rinchik and Russell, 1990).

Smith *et al.*, (1995), developed a strategy for chromosomal translocation in embryonic stem (ES) cells that relies on sequential gene targeting and Cre-loxP site-specific recombination. Gene targeting was first used to integrate loxP sites at a desired c-*myc* and immunoglobulin heavy chain locus on chromosomes 15 and 12 respectively. A transient Cre-expressing plasmid was used to mediate site-specific recombination between the two chromosomal loci. Such a strategy will allow the design of a variety of chromosomal rearrangements that can be selected and verified in ES cells or activated in ES cell-derived mice for genetic studies.

That defined deficiencies, inversions and duplication extending to more than 1 cM (centi-Morgan) can be constructed in embryonic stem cells was shown by Ramirez-Solis *et al.*, (1995) using the Cre-loxP system. This was achieved by consecutive targeting of loxP recombination sites to the end points of a genetic interval followed by Cre induced recombination. The availability of mice with defined regions of segmental haploidy will allow their use in genetic screens and enable accurate models of human chromosomal diseases to be generated.

1.4.4 Other applications of the Cre-loxP site-specific recombination system

Antibody fragments comprising paired heavy (VH) and light (VL) chain variable domains can be displayed on the surface of filamentous bacteriophage and rare phage (encoding antigen binding activities) selected by binding to antigen (Phage display; McCafferty *et al.*, 1990). However, the heavy and light chain genes cannot be packaged together within the same phage particle, and so cannot not be simultaneously co-selected.

Waterhouse *et al.*, (1993) used the Cre-loxP system to combine the heavy and light chain genes from two different replicons within an infected bacterium. Two vectors were constructed; one encoding the light chain of a first antibody (and the heavy chain from a second, different antibody); and the other vector encoding the heavy chain of the first antibody. In both vectors the heavy (VH) chain variable domains were flanked by two loxP sites. In order to prevent the excision of the VH genes in the presence of Cre recombinase, one of the loxP sites corresponded to the wild-type sequence and the other contained a point mutation within the spacer region. The wild-type loxP and the mutant loxP site do not recombine with each other in the same vector but recombine with sites of matching sequence in different vectors. When Cre recombinase was provided *in vivo*, the two vectors co-integrated by recombination

between either mutant or wild-type loxP sites to create two other chimaeric antibodies. This strategy should allow the construction of extremely large combinatorial repertoires of antibody fragments for the rapid screening and expression of highly specific antibodies.

The isolation and analysis of genomic DNA sequences has been facilitated by the development of phage λ replacement vectors that can contain up to 23 Kb of contiguous DNA and cosmid vectors that can contain inserts of up to 40 Kb of DNA. Genomic libraries in phage λ are easy to screen (Sambrook *et al.*, 1989) but present several problems in the analysis of insert DNA and subcloning. Subcloning and analysis each require the preparation of the recombinant λ clone DNA. Holt and May (1993) developed a novel phage λ replacement vector that does not necessitate the preparation of phage λ DNA for subcloning and analysis of inserts, because the inserts contained in this vector can be converted to plasmids using the Cre-loxP system as described previously for automatic subcloning from cDNA cloning vectors (Maruyama and Brenner, 1992; Palazzolo *et al.*, 1990). The subcloning technique described in Maruyama and Brenner, (1992) uses a bacteriophage λ cloning vector which has a unique *Bam*HI site engineered into λ gam gene. Cloning of foreign DNA into this site allows selection of recombinant vector on bacteriophage P2 lysogens of *E. coli*. Furthermore, when the recombinant phage infect a Cre-producing *E. coli* strain, a site-specific recombination event between loxP sites on the cloning vector results in the excision of a plasmid replicon with the cloned insert. In addition single-stranded DNAs can be recovered by growing helper M13 phages on bacteria containing such plasmids. This λ cloning vector can be used instead of λZAP (Short *et al.*, 1988) because there is no direct selection for recombinants using λZAP or other selective phage λ vectors (e.g. pλZd39; Murphy and Schimke; 1991) which does not permit the excision of the plasmid *in vivo*.

The results achieved by various groups thus far, illustrate that the Cre-loxP site-specific recombination (and other site-specific recombination systems) can be useful as tools for the large scale DNA manipulation *in-vitro* and *in-vivo* of eukaryotic chromosomes.

1.5 Disadvantage(s) of the Cre-loxP recombination system

The main drawback of the Cre-loxP system for genomic engineering is that the recombination event is freely reversible because recombination occurs with base-pair precision. Hence active sites remain after the reaction has occurred. In site-specific DNA integration, recombination between a site on a circular molecule and a site on a chromosome results in the insertion of the circular DNA into the chromosome. However, the integrated DNA is flanked by two recombination sites of the same orientation and can therefore excise if the recombinase is present. Since intramolecular excision is kinetically favoured over bi-molecular integration, insertion products are inherently unstable in the presence of the recombinase.

For a co-integration recombination reaction, it may be necessary to limit the recombinase activity to prevent integrated molecule from being re-excised. This has been achieved by transient provision of the recombinase: either from an inducible promoter (Sauer, 1987; Sauer and Henderson, 1988), introducing a non-selected plasmid that expresses the recombinase transiently and is subsequently lost (Sauer and Henderson, 1990; Fukushige and Sauer, 1992) or even using purified recombinase enzyme which provides a transient source of activity (Baubonis and Sauer, 1993). Another strategy has been to introduce selective base changes into recombination sites such that the product sites of a recombination reaction are less likely to undergo further recombination (Albert *et al.*, 1995; described in chapter 5).

Despite this drawback, the Cre-loxP site-specific recombination system continues to be used as a tool for the large scale DNA manipulation both *in vivo* and *in vitro*.

1.6 Aims of the project

The objective of this work was two fold; the first was to develop a mammalian expression system incorporating the Cre-loxP site-specific recombination system as a tool that would allow the targeted integration of transgenes to pre-selected chromosomal regions in order to obtain reproducible expression levels of heterologous protein. The second involves the fundamental investigation of the Cre-loxP reaction *in vitro* with the aim of identifying irreversible substrates that could be used to improve the efficiency of targeted integration.

The advantages and disadvantages of mammalian cells for the expression of recombinant proteins over other expression systems (such as bacteria and lower eukaryotic systems) has previously been described. Owing to the variable expression levels of recombinant proteins observed in mammalian systems, various approaches to optimise expression has been investigated (described above), however only the Cre-loxP site-specific recombination system provides a strategy that overcome variability due to chromosomal position effects and differing copy numbers.

The first objective describes the use of the Cre-loxP site-specific recombination system as a tool for gene targeting in CHO DHFR$^-$ (Chinese hamster ovary cells which are deficient for dihydrofolate reductase gene) cell line. Initially a loxP site was targeted into the CHO genome via a plasmid encoding a reporter gene, in this case SAP (secreted alkaline phosphatase gene), and a selectable marker gene, *dhfr*. This plasmid will integrate randomly into the CHO chromosomes. Given that integration is random, this could occur either at a promoter region, an enhancer/repressor region, a non-coding region or a region in the chromosome under cell-cycle/tissue specific induction. Integration into any of these regions would have a direct effect on the level(s) of gene expressions; in this case the reporter gene, SAP. The level of SAP therefore allowed the selection of transformants in which the loxP site has been placed in a favourable region for gene expression.

The Cre-mediated site-specific targeting event was investigated by transfecting either a promoterless *lacZ*/neo^r fusion gene or a promoterless CAT reporter gene (each encoding a loxP site) into the chromosomal loxP site (Figure 1.3). A source of Cre recombinase was supplied by either co-transfecting with the targeting plasmid, purified Cre protein (purified from baculovirus/insect expression system, described below) or a *cre* expression plasmid. Proper targeting activates the expression of the promoterless *lacZ*/neo^r fusion gene or the promoterless CAT gene from the chromosomally placed β-actin promoter. CHO cells containing this target are resistant to the neomycin analogue G418 and show β-galactosidase or CAT activity. This targeting event was investigated in independent CHO cell lines each containing single copy insertions of the loxP site in a different region of the chromosome (indicated by the level of SAP activity). The expression level of the second reporter and the frequency of the targeting event from different cell lines were compared to establish the effect of the chromosomal environment on the efficiency of recombination and gene expression. The Cre-mediated targeted integration into CHO cells is described in chapter 5.

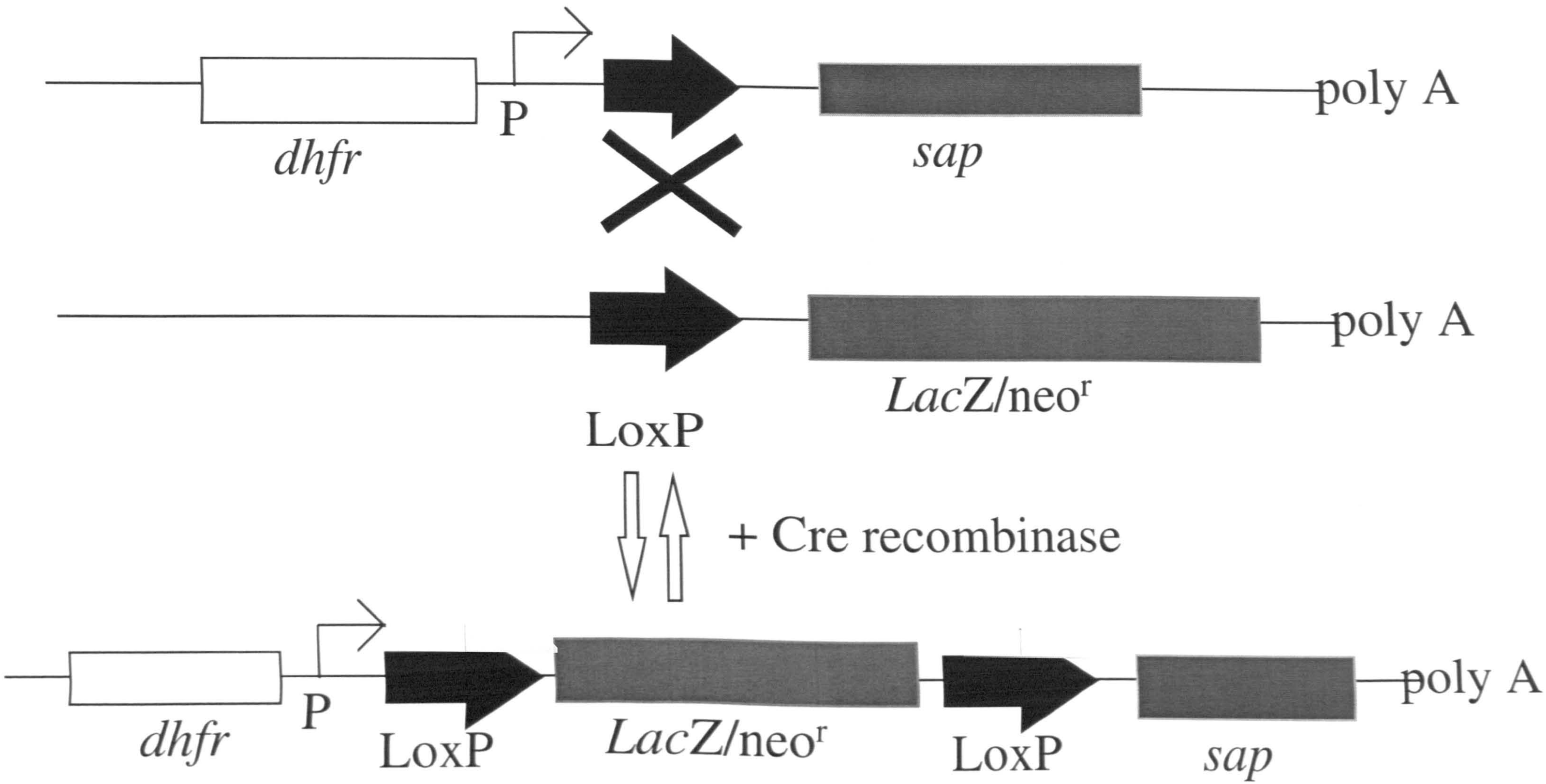


Figure 1.3-Schematic representation of Cre-mediated targeting of an exogenous DNA to a chromosomal loxP site. Initially a loxP site was placed in the genome of CHO DHFR$^-$ cells via a plasmid encoding a dihydrofolate reductase gene (*dhfr*) for selection and a secreted alkaline phosphatase (*sap*) reporter gene under the transcriptional control of β-actin promoter (P). Cre recombinase mediated targeting of a promoterless *lac*Z/neo^r fusion to the chromosomal loxP site activates the expression of the fusion gene from the β-actin promoter initially driving the SAP expression. This results in targeted clones which are resistant to the neomycin analogue G418, positive for β-galactosidase expression and give loss or reduction in SAP expression.

To address the second objective (and part of the first objective, which was the direct introduction of Cre protein into cells), the Cre protein was expressed in the baculovirus/insect expression system. Chapter 3 describes the construction of a recombinant baculovirus vector encoding the *cre* gene for expression in insect systems, a system that has been successfully used for the expression of high levels of active eukaryotic and prokaryotic proteins (Luckow and Summers, 1988). Two *in vitro* assay systems were set up to investigate the activity of the purified Cre protein: (i) A gel retardation assay was used to investigate the binding interaction of the Cre protein with oligonucleotide sequence corresponding to loxP sites. (ii) The ability of the purified Cre protein to mediate site-specific recombination *in vitro* was investigated. A loxP substrate for the *in vitro* analysis was constructed. This construct was designed such that loxP sites flanked on the one side an antibiotic resistance gene and on the other side a reporter *lac*Z gene. Cre-mediated recombination at the loxP sites will lead to the excision of the *lac*Z gene such that the recombination event could be analysed quantitatively either directly by restriction endonuclease digestion or by the phenotypic colour selection of substrates and products following an *in vivo* assay. The Cre-loxP *in vitro* assay system allowed the investigation of various kinetic parameters of the reaction and the manipulation of the loxP sites to achieve irreversible loxP substrates which could be used to improve the efficiency of *in vivo* targeted integration. Cre-mediated intermolecular integration event was also investigated using DNA fragments each containing a loxP site. The *in vitro* kinetic studies of the Cre-loxP reaction is described in chapter 4.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

Chemicals were purchased from Sigma or BDH, unless stated otherwise, and were Analytical Reagent quality (Analar) or of the highest grade available.

2.2 General DNA and sub-cloning techniques

2.2.1 Restriction endonuclease digestion of DNA

Restriction digests were done in the reaction conditions recommended by the enzyme manufacturer (Boehringer Mannheim; see Appendix 3 for buffer composition) using ~10 units of restriction enzyme/μg DNA, in a reaction volume at least 10 times the volume of enzyme used.

2.2.2 Ethanol precipitation of DNA from solution

DNA was concentrated or recovered from solution by ethanol precipitation. This was done by adding 0.1 volumes of 3 M sodium acetate pH 5.2 and 2.5 volumes of ice-chilled ethanol (99.5%) to the DNA-containing solution and mixed. The precipitated DNA was pelleted by centrifugation at 14000 rpm (15800 g) in a 5415 C Eppendorf centrifuge for 10 minutes at room temperature and washed twice with ice cold 70% ethanol. The DNA pellet was allowed to air dry and dissolved in an appropriate volume of TE (10 mM Tris-HCl, 1 mM EDTA pH 8.0) buffer.

2.2.3 Extraction of a DNA solution with organic solvent

DNA in solution was separated from protein contamination by extraction with an equal volume of phenol and chloroform mixed in the ratio 1:1 (chloroform contains 4% isoamyl alcohol, hereafter just referred to as chloroform). The two phases were mixed thoroughly by vortexing. The organic and aqueous phases were separated by centrifugation at 14000 rpm for 10 minutes. The upper aqueous phase containing the DNA was recovered avoiding any denatured protein at the interface. DNA solutions were generally extracted twice with phenol/chloroform.

2.2.4 Agarose gel electrophoresis of DNA

DNA fragments were separated according to size by horizontal electrophoresis through an appropriate percentage (0.5-1.2%) agarose gel made up in either TAE (40 mM Tris-acetate, 1 mM EDTA) or TBE (90 mM Tris-HCl, 90 mM boric acid, 2.5 mM EDTA, pH 8.0). Ethidium bromide (EtBr) was added at 0.5 µg/ml. Samples were mixed with 0.2 volumes of loading buffer (25% glycerol, 50 mM EDTA, 0.5% bromophenol blue) and electrophoresed in the same buffer as used to cast the gel at 1-10V/cm. The lambda (λ) Hind III marker, 100 bp ladder or the 1 Kb DNA ladder (GIBCO BRL; fragment sizes in Appendix 5-7) were used as the molecular weight size markers.

2.2.5 Recovery of DNA from an agarose gel

DNA fragments were recovered from agarose gel using the Prep-A-gene kit (Biorad), according to the manufacturer's protocol. In this method, the DNA band was sliced from the agarose gel and dissolved in sodium iodide solution. A silica matrix was then added to bind the DNA and DNA subsequently eluted from the silica matrix in high salt buffer.

2.2.6 Conversion of DNA fragments with 5' overhangs to blunt-ended fragments using Klenow

DNA fragments with 5' protruding ends were converted to blunt-ended fragments in a DNA synthesis reaction. The Klenow fragment of *Escherichia coli* (*E. coli*) DNA

polymerase 1 (Promega) was added at 5 units/µg DNA in a reaction buffer containing Mg^{2+} (e.g. Boehringher Mannheim buffer B; see Appendix 3 for composition). Deoxyribonucleotides (dNTPs) were added in excess (each at 25 µM) and the reaction incubated at room temperature for 30 minutes. The reaction was stopped by heat inactivating the enzyme at 70°C for 10 minutes and blunt-ended DNA fragments recovered by ethanol precipitation.

2.2.7 Conversion of DNA fragments with 3' overhangs to blunt-ended fragments using T4 DNA polymerase

DNA fragments with 3' overhangs were converted to blunt-ended fragments in a DNA synthesis reaction. T4 DNA polymerase (New England Biolabs) was added at 3 units/ µg DNA to a reaction containing Boehringer Mannheim restriction buffer A (see Appendix 3 for composition). Deoxyribonucleotides (dNTPs) were added in excess (each at 25 µM) and the reaction incubated at room temperature for 30 minutes. The reaction was stopped by the addition of EDTA pH 8 to a final concentration of 20 mM and blunt-ended DNA fragments recovered by ethanol precipitation.

2.2.8 S1 nuclease digestion

DNA fragments with sticky ends were hydrolysed with 0.1 units of S1 nuclease (GIBCO BRL)/µg DNA in a 500 µl reaction volume containing S1 nuclease buffer. The reaction was done at 37°C for 10 minutes. At 0.5, 1, 5, and 10 minute time points, 100 µl of the reaction mixture was removed and the reaction stopped by the addition of EDTA pH 8 to a final concentration of 50 mM. The hydrolysed DNA fragments were recovered by ethanol precipitation.

2.2.9 Dephosphorylation of linear DNA fragments using calf intestinal alkaline phosphatase

Removal of 5' phosphate residues from DNA fragments were done using calf intestinal alkaline phosphatase (CIP, Promega). The enzyme works on both blunt and sticky ends of DNA fragments. The reaction was done in Boehringer Mannheim restriction buffer M (see Appendix 3 for composition; although other Boehringer buffers work as well), by the addition of 2 units CIP/µg DNA followed by incubation

at 37°C for 30 minutes. The reaction was stopped by heat inactivation at 70°C for 10 minutes and the dephosphorylated fragments purified by phenol/chloroform extraction followed by ethanol precipitation.

2.2.10 Phosphorylation of linear DNA fragments using bacteriophage T4 polymerase kinase

The 5' termini of double-stranded oligonucleotide sequences were phosphorylated by T4 polymerase kinase (Promega) using 5 units/µg DNA in a reaction volume containing the kinase buffer (Promega). The reaction was incubated at 37°C for 1 hour and stopped by heat inactivation at 70°C for 10 minutes. The phosphorylated DNA fragments were then recovered by ethanol precipitation.

2.2.11 Ligation of DNA fragments

Vector and insert fragments were mixed in a molar ratio of 1:3 to give ~100 ng DNA in a final volume of 16 µl. To this, 2 µl 10 × ligase buffer (Promega) and 2 µl T4 DNA ligase enzyme (Promega) were added and the mixture incubated at 17°C overnight. Approximately half of this ligation reaction was used in transformation reactions in *E. coli* DH5α competent cells (described in section 2.2.13).

2.2.12 Preparation of *E. coli* DH5α competent cells

Bacteria were made competent for transformation by plasmid DNA following the method of Hanahan (1983). SOB medium (2% Bacto tryptone. 0.5% Bacto yeast extract, 10 mM NaCl, 25 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) of which 10 ml was inoculated with a single colony of *E. coli* DH5α and grown overnight, shaking at 37°C. SOB medium (500 ml) was then inoculated with this overnight culture and grown at 37°C to A_{550} 0.4-0.6 (4-7 × 10^7 viable cells/ml). The culture was chilled on ice for 30 minutes and cells pelleted by centrifugation at 3000 rpm (1465 g) in a Sorval GSA rotor for 15 minutes. The cells were resuspended in 200 ml of ice-cold TFB1 solution (30 mM KOAc, 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MgCl_2$, 15% Glycerol, adjusted to pH 5.8 with acetic acid and filter sterilised). The cells were pelleted as before and resuspended in 20 ml ice-cold TFB2 solution (10 mM 3-[-morpholino] propanesulfonic acid {MOPS}, 75 mM $CaCl_2$, 10 mM $RbCl_2$, 15%

Glycerol, adjusted to pH 6.5 with KOH and filter sterilised). Aliquots (200 µl) were snap frozen by transferring to microcentrifuge tubes in a dry ice bath and then stored at -70°C.

2.2.13 Transformation of *E. coli* DH5α competent cells

All plasmids used contained the gene encoding β-lactamase. Transformation was done following the method of Hanahan (1983). Approximately 50 ng of a ligation reaction or 10 ng of purified plasmid DNA was added to 200 µl of thawed DH5α competent cells and incubated on ice for 5 minutes. Cells were heat shocked for 2 minutes in a 42°C water bath, and returned to ice for 5 minutes. After the heat shock treatment, 800 µl of SOB medium was added to the transformed cells and then incubated at 37°C for 1 hour. Transformed cells (200 µl) were then spread onto selective agar plates (LB [10 g Bacto-trypton, 5 g Bacto yeast extract, 10 g NaCl in 1L], 1.5% agar + 100 µg/ml ampicillin) and incubated at 37°C overnight. Transformed cells where the plasmid genotype (resistance to ampicillin) was expressed, were observed the next day as white colonies.

2.2.14 Small-scale plasmid preparation

The method used is a modification of the alkaline lysis method of Birnboim and Doly (1979). A single transformed bacterial colony was picked into 3 ml LB medium and grown (to stationary phase) overnight with shaking at 37°C. The culture (2 ml) was transferred to an Eppendorf tube and centrifuged at 14000 rpm (15800 g) for 5 minutes and the cell pellet resuspended in 200 µl of ice-cold 50 mM glucose, 25 mM Tris-HCl, pH 7.5, 1 mM EDTA solution. Cells were lysed by the addition of 200 µl 0.2 M NaOH, 1% SDS and incubated on ice for 5 minutes. Neutralisation solution (2.55 M sodium acetate, pH 4.8) of which 200 µl was then added to the cell extract and incubated for further 5 minutes on ice. Precipitated chromosomal DNA, protein and SDS was pelleted by centrifugation at 14000 rpm (15800 g) for 10 minutes. The supernatant was transferred to a fresh tube and the plasmid DNA precipitated by the addition of 1 ml of 95% EtOH. The DNA was pelleted by centrifugation at 14000 rpm (15800 g) for 10 minutes, washed with 70% ethanol, dried and the pellet redissolved in 50 µl TE buffer pH 8, containing DNAase free pancreatic RNAase (50 µg/ml) and stored at -20°C.

For some purposes, plasmid DNA was extracted using the Promega Magic/Wizard mini-prep kit following the manufacturers instruction. This is based on an alkaline lysis method and utilises a DNA-binding resin.

2.2.15 Large-scale plasmid preparation

The alkaline lysis method of Birnboim and Doly (1979), was followed for the large scale preparation of plasmid DNA. A single transformed bacterial colony was picked into 10 ml LB medium and grown (to stationary phase) overnight with shaking at 37°C. LB medium (500 ml) was inoculated with the 10 ml overnight culture and grown overnight with shaking at 37°C. Bacteria cells were pelleted by centrifugation at 5000 rpm (4080 g), in a Sorvall GSA rotor for 10 minutes and resuspended in 25 ml of 25 mM Tris-HCl, pH 8, 10 mM EDTA, 50 mM glucose. Cells were lysed by the addition of 50 ml of 0.2 M NaOH, 1% SDS, with gentle mixing. Chromosomal DNA, protein and SDS were precipitated by the addition of 35 ml 2.5 M potassium acetate, 2.5 M glacial acetic acid followed by incubation on ice for 10 minutes. The solution was centrifuged at 8000 rpm (10400 g) in a Sorvall GSA rotor, at 5°C for 15 minutes and the supernatant recovered and filtered through 3 layers of gauze to remove any precipitate. Plasmid DNA was precipitated by the addition of 70 ml isopropanol and pelleted by centrifugation at 8000 rpm for 10 minutes at room temperature. The DNA was dissolved in 9 ml TE buffer, pH 8. This crude plasmid DNA solution was transferred to a Corex centrifuge tube containing 10.5 g of CsCl. Tris-HCl, pH 8 (1 ml), 120 µl of 0.5 M EDTA, pH 8 and 320 µl of EtBr (10 mg/ml). The solution was clarified by centrifugation (HB4 rotor) at 8000 rpm for 15 minutes at room temperature. The clear supernatant was transferred to a Beckman Quick-seal tube and centrifuged at 45000 rpm at 22°C for 40 hours (Beckman ultra-centrifuge 70.1. Ti rotor). This generated an equilibrium gradient in which different DNA species migrated as distinct band to different heights corresponding to their buoyant densities. Two bands of DNA were visible under normal light; an upper band of nicked plasmid DNA and a denser lower band of covalently closed circular plasmid DNA. The lower band was collected using a hyperdermic-needle connected to a 2 ml disposable syringe pushed through the side of the tube. The top of the tube was punctured before the needle was gently introduced into the lower band.

Ethidium bromide bound to the DNA was removed from the plasmid solution (volume noted) by repeated extractions with water-saturated butan-1-ol until the upper phase solution was colourless. The solution containing the plasmid DNA was

transferred to a Corex centrifuge tube, the volume was made up to 4 times that noted above with sterile water, and two volumes of 95% ethanol added. Precipitated plasmid DNA was recovered by centrifugation at 8000 rpm in a Sorvall SS34 rotor (7710 g) for 30 minutes. The DNA was washed twice with 70% ethanol and then dissolved in 500µl of TE, pH 8. The concentration and purity of the DNA solution were determined spectrophotometrically as described in section 2.2.17.

2.2.16 Genomic DNA preparation

The method of Laird *et al.*, (1991), for isolating mammalian DNA was followed with some modification. Mammalian cells growing on a 6-well plate or pelleted by centrifugation, were washed with PBS and lysed by the addition of between 0.5 to 1 ml lysis buffer (100 mM Tris-HCl, pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 µg/ml Proteinase K). Lysis was done at room temperature for a minimum of 6 hours with agitation. One volume of isopropanol was added to the lysate and the samples mixed by swirling until precipitation was complete. The precipitated DNA (if lysis was done on a 6-well plate) solution was carefully transferred to a microfuge tube and centrifuged at 14000 rpm (10508 g) for 10 minutes. The liquid was discarded and precipitated DNA washed once with 70% ethanol. The DNA was dried at room temperature and resuspended in between 50 to 500 µl of 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.5.

2.2.17 Determination of DNA concentration

The A_{260} and A_{280} of an aliquot of DNA solution diluted 1 in 100, were measured. Given that an A_{260} of 1 corresponds to 50µg/ml of double-stranded DNA solution, the concentration of the DNA solution was calculated. Pure DNA solutions have an $A_{260/280}$ ratio of 1.8, lower $A_{260/280}$ than 1.8 indicates protein contamination.

DNA samples of very low concentration or low molecular weights were estimated by electrophoresing an aliquot (known volume) on an agarose gel with a dilution series of linearised plasmid standards of known concentrations. Comparison of the intensity of the sample band with the DNA concentration standards under UV illumination was used to estimate the concentration of DNA.

2.2.18 Southern blotting

Localisation of particular sequences within plasmid or genomic DNA was determined by the transfer techniques described by Southern (1975), and Meinkoth and Wahl (1984). DNA fragments, separated by electrophoresis on an agarose gel, were transferred to a nylon membrane prior to hybridisation with radiolabelled probes.

After electrophoresis, DNA in the gel was denatured and slightly fragmented by soaking the gel in 0.25 M HCl for 10 minutes, followed by treating with 1.5 M NaCl, 0.5 M NaOH for 30 minutes with one change of the solution. The gel was then neutralised by soaking in 3 M NaCl and 0.5 M Tris-HCl pH 8.0 for two 30 minutes washes. Denatured DNA fragments were then transferred to a nylon membrane. The transfer apparatus consisted of a tray containing 20 × SSC over which was placed a glass plate slightly narrower than the tray; two sheets of Whatman 3 MM paper soaked in 20 × SSC were placed on top of the glass plate with their ends dipping into the 20 × SSC solution in the tray. The gel was placed directly onto the paper making sure that no air bubbles were trapped underneath. A piece of nylon filter membrane (Hybond-N, Amersham LIFE SCIENCES) pre-cut to the size of the gel was soaked in water for 15 minutes, washed briefly in 20 × SSC and then placed onto the gel taking care to avoid air bubbles. Four sheets of Whatman 3 MM paper cut to the size of the gel were soaked in 20 × SSC and placed on top of the membrane and surrounded by saran-wrap, followed by a stack of paper towels 8-10 cm high. The various layers were compressed using another glass plate and a 1.5 Kg weight. Transfer was allowed to proceed overnight. The apparatus was dismantled, nylon membrane separated from gel and rinsed briefly in 2 × SSC to remove any adherent agarose particles. The filter was then air dried and DNA fixed onto the membrane using a U.V crosslinker.

2.2.19 Preparation of radioactively labelled DNA probes

DNA fragments were labelled to high specific activity with ^{32}P using the random primer kit (Stratagene), following the manufacturers protocol. The method relies on the ability of random hexanucleotides to anneal to multiple sites along the length of a DNA template. The primer-template complex is then a substrate for Klenow fragment of DNA polymerase I. The enzyme synthesises new DNA by incorporating nucleotide monophosphates at the free 3'-OH group provided by the primer. The newly synthesised DNA is made radioactive by substituting a radiolabelled nucleotide for a non-radioactive one in the reaction mixture.

2.2.20 Hybridisation of Southern blot

Southern blots (described in section 2.2.18) were hybridised with radioactively labelled DNA probes. Prehybridisation and hybridisation of Southern blots were done in a hybridisation buffer. The hybridisation buffer consists of 21.3 ml of 100 × Denhardt's solution (2g Ficoll 400, 2g polyvinylpyrolidine [PVP], 2g BSA fraction V in 100 ml [Denhardt, 1966]), 22.4 g NaCl, 11.2 g Na_2Citrate, 42.6 g Dextran sulphate (Pharmacia) and 0.5% SDS made up to 500 ml.

Just before pre-hybridisation, 0.5 ml of denatured salmon sperm DNA (10 mg/ml, GIBCO BRL), was added to 10 ml of the hybridisation buffer at 65°C. The mixture was then added to the Southern blot presoaked in 2 × SSC in a hybridisation flask. Prehybridisation was done at 65°C for 4 hours in a Hybaid hybridisation oven. The pre-hybridisation solution was subsequently discarded and 10 ml of pre-warmed (at 65°C) hybridisation buffer added to the flask. Heat denatured radiolabelled DNA probe (by incubating the radiolabelled DNA probe at 100°C for 5 minutes) and 0.5 ml of denatured salmon sperm DNA were added to the Southern blot and hybridised overnight. The probe solution was discarded by pouring it down a designated radioactive disposal sink. The filter was washed twice for 20 minutes with 2 × SSC and 0.1% SDS at 65°C. The filter was repeatedly washed in a higher stringency solution of 0.1 × SSC and 0.1% SDS at 65°C for 30 minutes, until most of the non-specific probe-DNA interactions were removed, using Geiger monitor to detect background. The filter was then wrapped in Saran wrap, removing excess liquid but keeping the filter damp. Hybridisation signals were detected by autoradiography.

2.2.21 DNA Sequencing

Sequencing of plasmid DNA was done by the dideoxy chain termination method (Sanger *et al.*, 1977) using sequenase kit Version 2.0 (United States Biochemicals), following the protocol supplied with some modifications. Plasmid DNA preparation (1-5 µg DNA) in a volume of 18 µl was mixed with 2 µl of NaOH/EDTA (2 M/2 mM), and incubated at 65°C for 15 minutes. The denatured DNA was precipitated by adding 2 µl of 3 M sodium acetate pH 5.2 and 50 µl of ethanol (95%), mixed and incubated on dry ice for 10 minutes. The precipitated DNA was pelleted by centrifugation at 14000 rpm for 10 minutes and washed with cold 70% ethanol. The DNA was resuspended in 7 µl of distilled water and mixed with 2 µl of 5 × sequencing buffer (0.5 M Tris-HCl, pH 8, 10 mM $MgCl_2$, 150 mM NaCl), and 1 µl of

primer (0.2 µg/µl). The mix was incubated at 65°C for 15 minutes and then cooled to room temperature. The annealed plasmid DNA and primers were used in labelling and termination reactions as described in the kit protocol, using [α^{35}S]-dATP at > 1000 Ci/mmol (Amersham).

An 8% polyacrylamide/urea gel was made by adding 400 µl 10% ammonium persulphate (AMPS; BDH) to 75 ml acrylamide + urea solution (Acryl-a-mix 8, Promega; i.e. acrylamide to bis-acrylamide at 19:1, 8M urea, 1 × TBE, 0.05% N,N,N',N', tetramethylethylenediamine [TEMED]), to initiate polymerisation. Immediately before loading, the reactions were denatured by heating to 95% for 5 minutes. The sequencing reaction was analysed by electrophoresis through the 8% polyacrylamide/urea gel in TBE buffer for 2 hours at 70W. Gels were then fixed in 5% methanol, 10% acetic acid for 10 minutes and then transferred to a Whatman 3 MM paper and dried under vacuum at 80°C. Signals were visualised by autoradiography.

2.2.22 General polymerase chain reaction (PCR)

The PCR technique used is a modification of Saiki *et al.*, (1988). A pair of DNA oligonucleotide primers (0.2µg of each) were used to amplify DNA sequences from a plasmid (0.1µg) or genomic (0.1µg) template DNA in a PCR reaction. The reactions were done in a final 50µl volume containing 5 µl 10 × *Taq* polymerase buffer (Boehringer Mannheim), 1 µl each of dATP, dTTP, dCTP and dGTP (of a 10 mM stock each) and 0.5µl *Taq* polymerase (Boehringer Mannheim). The reaction was overlaid with mineral oil and placed in a thermocycler. The PCR reactions were run for 35 cycles (30 seconds at 95°C, 2 minutes at 60°C, 3 minutes at 72°C) unless otherwise stated. Aliquots (10µl) of the PCR products were separated on a 0.7% agarose gel and DNA visualised by ethidium bromide staining.

2.3 Mammalian cell culture techniques

2.3.1 Cell line

The cell line used is the Chinese Hamster Ovary (CHO-DUKX B11, American Type Culture Collection, Reference: CCL 61), obtained from the Wellcome Research Labs., Kent (now Glaxo Wellcome Medicines Centre, Stevenage).

2.3.2 Cell Culture Media

2.3.2.1 Complete Medium

The complete medium used for the maintenance of CHO-DUKX B11 cell line was Iscove's modified Dulbecco's medium (Iscove's DMEM, containing L-Glutamine, GIBCO BRL). The medium was supplemented with 10% foetal calf serum (FCS, Advanced Protein Products Ltd), non-essential amino acids (NEAA; alanine, aspartate, glycine, serine, proline, glutamate and asparagine each at 100 µM; Flow/ICN), penicillin (100 units/ml), streptomycin (100 µg/ml), and HT medium (0.1 mM hypoxanthine/0.01 mM thymidine, GIBCO BRL).

2.3.2.2 Select Medium

The medium used for the selection of dihydrofolate reductase positive ($DHFR^+$) phenotype was Iscove's DMEM as above supplemented with non-essential amino acids, penicillin (100 units/ml), streptomycin (100 µg/ml), as above, 10% dialysed FCS (Advanced Protein Products Ltd), and omitting the hypoxanthine/thymidine solution.

2.3.3 Maintenance of CHO-DUKX B11 cell line

CHO-DUKX B11 cells were grown in Iscove's-DMEM complete medium as adherent cultures, in flasks at 37°C in a 5% CO_2 atmosphere. Adherent cells were passaged when they reached confluency (~80% viability), approximately 3-4 days. Culture

medium was discarded and the cells washed with phosphate buffered saline (PBS, GIBCO BRL). Cells were detached from the flask wall by incubation with versene (0.2 g EDTA/L in an isotonically buffered saline; GIBCO BRL) + 0.25% trypsin (GIBCO BRL). Approximately 5 ml of the versene/trypsin solution was added to a T75 flask and left for 3 minutes at room temperature. Following this, the dislodged cells were diluted with 5 ml of fresh medium. Cell suspension was transferred to a 15 ml Falcon centrifuge tube (Becton Dickinson), and pelleted by centrifugation in a Sorvall RT6000B centrifuge (as were all further mammalian suspensions), at 1000 rpm (~ 135 g) for 5 minutes. The cell pellet was resuspended in 5 ml of fresh medium, an aliquot was diluted ten fold with trypan blue solution (0.4% in saline, ICN Flow), and viable cells (non-stained) counted using a haemocytometer. Tissue culture flasks were seeded at ~ 2×10^5 cells/ml.

2.3.4 Freezing cell lines for long term storage in liquid nitrogen

CHO-DUKX B11 cells in exponentially growing phase (~80% viability) were pelleted by centrifugation as above. The cells were resuspended in ice-cold freezing medium (i.e. complete culture medium containing 20% FCS and 10% dimethyl sulphoxide [DMSO]), at ~ 4×10^6 cells/ml. Aliquots (1 ml) were transferred to cryovials and placed in Bicell freezing vessels, in a -70°C freezer overnight, and then transferred to liquid N_2 the following day.

2.3.5 Reviving frozen cells

Vials of cells were removed from liquid N_2 and the contents thawed immediately. Cells were slowly added to 9 ml of pre-warmed medium in a 15 ml Falcon tube, to allow the DMSO to diffuse out of the cell membrane. The cells were pelleted by centrifugation as above, and the supernatant, containing DMSO, discarded. Cells were resuspended in 12 ml of medium and transferred to a T75 flask.

2.4 Mammalian cell transfection techniques

2.4.1 Transient transfection of CHO-DUKX B11 cells using Transfectam

Transfectam, a synthetic cationic lipopolyamine molecule (Promega), was used following the manufacturer's instructions with optimisation for CHO-DUKX B11 cells. Transient transfections were done in a 6-well tissue culture plate. Wells were seeded with 2.5×10^5 cells, a day before transfection. For each well of cells to be transfected, 20µg Transfectam, DNA (different amount) and 0.5 ml serum-free medium (SFM) were mixed together in a polystyrene Bijoux. This mix was incubated for 10 minutes at room temperature, to allow transfectam and DNA complex to form. Meanwhile, adherent cells (approximately 75% confluent), were washed twice with SFM (serum interferes with the action of transfectam). The DNA/transfectam/SFM mix and a further 0.5 ml SFM were added to each well after the second wash and the plates incubated at 37°C for 4 hours. The transfection mix was removed and each well fed with 1 ml serum-containing Iscove's DMEM medium. The plate was returned to the 37°C incubator for 1 to 2 days. Transfected cells were harvested by centrifugation and plasmid DNA isolated using Promega mini-prep kit.

2.4.2 Stable transfection of CHO-B11 cells using Transfectam

The method described for transient transfection in section 2.4.1 was followed, but scaled up fourfold. A T25 flask was seeded with 1×10^6 cells the day before transfection. Plasmid DNA (16 µg) was mixed with 80µg Transfectam and the mixture added to the CHO-DUKX B11 cells and incubated at 37°C for 4 hours. Cells were fed with 15 ml of complete serum medium and incubated without selection for 2 days. Transfected cells were dilution cloned into select medium as described in section 2.4.3.

2.4.3 Dilution cloning of stably transfected CHO-DUKX B11 cells

Transfected cells growing in complete medium were harvested 2 days after transfection and counted. Cells were suspended in select medium and viable density adjusted to 1.5×10^3 cells/ml and diluted in select medium to 500 cell/ml, 150 cell/ml, 50 cell/ml, 15 cell/ml, 5 cell/ml and 1.5 cell/ml. The diluted cells were plated

out in 96 well plates at 200 µl/well to give 100, 30, 10, 3, 1, and 0.3 cells/well. The plates were incubated at 37°C for approximately three weeks until individual clones appeared. The supernatant from wells containing single clones was assayed for secreted alkaline phosphatase (SAP; see section 2.8.3) expression and positive clones expanded. Clones were named according to the plate and well co-ordinates from which they were picked.

2.4.4 Targeted transfection of CHO-DUKX B11 transformants using LipofectACE

LipofectACE reagent, a 1:2.5 (w/w) liposome formulation of dimethyl dioctadecylammonium bromide and dioleoyl phosphatidylethanolamine (Life Technologies, Inc.) was used essentially as suggested by the manufacturer with the following modifications. Cells were seeded in a six well tissue culture dishes and grown to ~80% confluency (1×10^5 cells). Plasmid DNA and recombinant Cre protein (purified from baculovirus-insect system), or DNA and pBS (Bluescript, Stratagene), were mixed with opti-MEM (Life Technologies Inc.) to give a final volume of 100 µl, and incubated at 0°C for 5 minutes. In a separate aliquot, 12 µl of the lipofectACE was added to 88 µl opti-MEM. The two aliquots were combined and incubated for 10 minutes at room temperature. The volume was brought to 1 ml by the addition of opti-MEM and the mixture added to a single well of cells that had been previously washed twice with opti-MEM. Cells were incubated for 5 hours at 37°C, 1 ml of $DHFR^+$ select medium was added and incubation continued overnight. The next day, the medium was replaced with fresh $DHFR^+$ select medium. Selection for targeted transformants was done 48 hours after the lipofection procedure in $DHFR^+$ select medium supplemented with G418 (400 µg/ml final concentration, Sigma). G418 resistant clones observed (2-3 weeks later), were picked and propagated for further analysis.

2.4.5 Stable transfection of CHO-DUKX B11 by electroporation

Exponentially growing cultures of CHO-DUKX B11 cells were harvested by versene + trypsin treatment and the viable cell density determined. For each electroporation, 1×10^7 cells in a 0.8 ml of complete medium were transferred to a sterile polystryrene Bijoux containing plasmid DNA. The cells/DNA mix was transferred to a sterile electroporation cuvette (Biorad, 0.4 cm electrode gap) at room temperature. Cells

were pulsed once at 250V, 960μF, using a Biorad gene pulser. The cuvette was returned to room temperature for 5 minutes, after which the contents were added to 12 ml of complete medium in a T75 flask, and returned to the 37°C incubator. Two days after transfection cells were harvested and counted. The cells were dilution cloned (as described in section 2.4.3) in DHFR$^+$ select medium and returned to 37°C for 2-3 weeks, until resistant clones appeared in the wells. Wells containing single clones were assayed for SAP expression and positive clones expanded.

Alternatively, cells for electrporation were washed with 10 ml of cold PBS after viable cell density had been determined. Cells were resuspended in 0.8 ml of cold PBS (to 1×10^7 cells), and added to a sterile polystyrene Bijoux containing DNA with or without Cre protein. The mix was transferred to a sterile electroporation cuvette and placed on ice for 5 minutes. Cells were pulsed once at 250V, 960μF using a BioRad gene pulser. The cuvette was returned to ice for 5 more minutes after which the cells were transferred to T75 flask and grown as described above.

2.4.6 Expanding CHO-DUKX B11 transformed clones

To expand single clones expressing SAP in 96-well plates, a few drops of versene + trypsin solution was added to each well. Dislodged cells were transferred to one well of a 24-well plate, containing 0.5 ml of fresh DHFR$^+$ select medium. Plates were placed in a 37°C incubator. When the cells reached confluency, they were assayed for SAP production, and, if still positive, they were expanded further from the 24-well plate to a 6-well plate, T25 and T75 flasks, using appropriate amount of versene + trypsin to detach the adhered cells at each stage. G418 resistant clones were expanded in DHFR$^+$ select medium containing G418 (400 μg/ml).

2.5 Insect cell and Baculovirus Techniques

2.5.1 Maintenance of SF9 cell line

The *Spodoptera frugiperda* cell line Sf9 (originated from IPLB-SF-9 cells, Vaughn *et al.*, 1977; obtained from WRL [now GWR&D]), was maintained in TC 100 medium (GIBCO BRL Life Technologies) which was supplemented with 10% foetal calf

serum (Advanced Protein Products), L-Glutamine (GIBCO BRL), and antibiotic-antimycotic solution (GIBCO BRL). All Sf9 cells were grown in this (complete) medium unless stated otherwise. For serum-free TC 100 medium, the FCS and antibiotic-antimycotic solutions were omitted. Cells were maintained as suspension cultures, grown in spinner flasks at 17°C. Flasks were seeded at approximately 0.3×10^6 cells/ml and passaged after 3 to 4 days, when the density had reached approximately 1.5×10^6 cell/ml.

2.5.2 Production of recombinant baculovirus

BacPAK6, an engineered *Autographa Californica* DNA digested with *Bsu*36I (Clontech; Fig 3.2), was used as the viral genomic DNA into which the foreign gene was targeted. Production of recombinant baculovirus was done essentially by following the manufacturers protocol. A 6-well plate was seeded with 2×10^6 Sf9 cells/well (one well for each recombinant virus to be made), and the cells allowed to adhere for 1 hour at 28°C. For each transfection, 100 ng of BacPAK6 DNA digested with *Bsu*36I, 500 ng of transfer vector containing the *cre* gene and 5 µg of lipofectin (GIBCO BRL) were mixed together in a polystyrene container and incubated at room temperature for 15 minutes to allow lipofectin-DNA complex to form. The adhered insect cells were washed twice with serum free TC 100 medium (lipofectin-mediated transfection is inhibited by a component of serum), and 1.5 ml of serum free medium added to each well. The Lipofectin-DNA mixture was added dropwise to a single well of cells and the plate incubated at 28°C for 5 hours to allow co-transfection of the cells with transfer vector and BacPAK6 DNA to take place. Subsequently, 1.5 ml of complete TC 100 medium was added to each well to stop the action of the lipofectin and incubation continued for 2-3 days at 28°C. Recombinant viruses formed as a result of homologous recombination between BacPAK6 and the transfer vector were released by cell lysis into the medium. The medium was harvested and used to infect Sf9 cells for plaque purification in order to isolate a single recombinant virus clone.

2.5.3 Plaque purification of recombinant baculovirus

A 6-well plate was seeded with 2×10^6 Sf9 cells/well and the cells allowed to adhere for 1 hour at 28°C. The supernatant from insect cells (co-transfected with BacPAK6 and the transfer vector) putatively containing recombinant virus was serially diluted in complete TC 100 medium to a million fold. The medium was removed from the cells

and replaced with 1 ml of the diluted virus solution per well. The plate was rocked for 1-2 hours at room temperature during infection. The inoculum was removed from each well and the infected cells overlaid with 2 ml agarose/TC 100 (1g of Seaplaque agarose autoclaved in 16 ml distilled water, cooled to 37°C and 32 ml warmed complete TC 100 medium added). The agarose was allowed to set for 15 minutes and then 2 ml of complete TC 100 was added to each well and plate returned to 28°C incubator for 2-3 days.

Plaques were visualised by staining the cells with 100µl Neutral Red/well for 2 hours and destained by inverting the plate at 4°C overnight. To isolate a clonal recombinant virus, a plug of agarose over an isolated plaque was removed, using a wide-bored plastic pastette and dropped into 1 ml of complete TC 100. Recombinant virus dispersed out of the plug into the medium after several hours of incubation at 4°C.

A second round of plaque purification was done to ensure that a recombinant virus was clonal by following the same procedure as above; by infecting cells with 10^{-1}, 10^{-2} and 10^{-3} dilutions of the medium containing recombinant virus dispersed from the agarose plug. A well separated plaque was picked and assumed to contain a clonal recombinant virus.

2.5.4 Preparation of a high titre virus stock

One well of a 6-well plate was seeded with 2×10^6 Sf9 cells and the cells allowed to adhere for 1 hour at 28°C. Complete TC 100 medium (1 ml) containing recombinant viruses (isolated by agarose plug purification) was used to infect the cells for 1-2 hours, rocking at room temperature. Complete TC 100 medium (1 ml) was then added to the well and the plate incubated at 28°C for 3-4 days, after which time the majority of the cells would have been infected and the replicated virus released into the medium. This high titre supernatant was harvested and passed through a 0.2µm filter unit (Coster) to remove cell debris. Larger-scale infections, following a similar protocol, gave rise to larger volumes of high titre, filtered virus stock.

2.5.5 Titration of recombinant virus

Using a method essentially the same as that for the plaque purification of recombinant virus (Section 2.5.3), high titre viral stocks were serially diluted in complete TC 100,

and 500 µl of the 10^{-4} to 10^{-9} dilutions used to infect 2×10^6 Sf9 cells/well in a 6-well plate. Infected cells were overlaid with agarose/TC 100 and 2 ml complete TC 100, and incubated at 28°C for 3-4 days. Following staining with neutral red and destaining, wells containing between 10 and 100 plaques were scored to give the number of plaque forming units (pfu) in 500 µl of a certain dilution of the virus. The titre, in pfu/ml, of the virus stock was calculated and confirmed by repeating the titration.

2.5.6 Infection of Sf9 cells

Small-scale infections of Sf9 cells were generally done in a 12-well plate. Cells were plated at 1×10^6/well and allowed to adhere for 1 hour. The medium was removed and replaced with approximately 0.5 ml complete TC 100 containing an aliquot of the viral stock, to give the required multiplicity of infection (designated time 0 hours post-infection). The plate was rocked at room temperature for 1-2 hours, during which time the cells were infected. A further 0.5 ml of medium was added to the wells and the plate incubated at 28°C until the supernatant and/or cells were harvested.

2.5.7 Recombinant baculovirus DNA preparation

Sf9 (2×10^6) cells were infected with the recombinant virus at 10 pfu/cell. After 2-3 days incubation at 28°C, the culture medium containing virus at high titre was harvested and filtered to remove cell debris. 1 ml of this clarified supernatant was transferred to a microcentrifuge tube and the virus preparation pelleted by centrifugation at 14000 rpm (15800 g) for 30 minutes at 4°C. The virus preparation was lysed by gentle resuspension in 100 µl of 10 mM Tris-HCl, pH 7.6, 10 mM EDTA, 0.25% SDS. The protein in the lysate was degraded by the addition of Proteinase K to a final concentration of 0.5 mg/ml and incubated at 37°C for 1 hour. The viral DNA in solution was isolated by extraction with an equal volume of phenol-chloroform. Extraction was done twice by gentle mixing of the two solutions (with no vortexing because the high molecular weight viral genomic DNA is susceptible to shearing forces). The upper aqueous phase containing the DNA was recovered avoiding any denatured protein at the interface. The DNA was precipitated by adding 0.1 volumes of 3 M sodium acetate pH 5.2 and 2.5 volumes of ice-chilled ethanol (95%) and mixed. The precipitated DNA was pelleted by centrifugation at 14000 rpm (15800 g) for 10 minutes at room temperature and washed twice with ice-cold 70%

ethanol. The DNA pellet was allowed to air dry and dissolved in 50 µl TE buffer pH 8.0 containing DNAase free pancreatic RNAase (50 µg/ml) and stored at -20°C.

2.5.8 Screening recombinant virus by the polymerase chain reaction (PCR)

A combination of oligonucleotide sequences designed as forward primer (BK59) to anneal within the polyhedrin locus and a reverse primer (CB1741; see appendix 2 for sequence) to anneal at the 5' end of the *cre* gene were used to amplify DNA sequences from the viral DNA. For the PCR reaction, 0.1 µg of viral DNA and 1 µl of each primer (0.2 µg/µl) were used. The PCR reaction was done in a final 100 µl volume containing 10 µl 10 × *Taq* polymerase buffer (Boehringer Mannheim), 2 µl each of dATP, dTTP, dCTP and dGTP (of a 10 mM stock each) and 0.5 µl *Taq* polymerase (Boehringer Mannheim). The reaction was overlaid with mineral oil and placed in a thermocycler. The PCR reaction was run for 35 cycles (30 seconds at 95°C, 2 minutes at 60°C, 3 minutes at 72°C). An aliquot (10 µl) of the PCR product was separated on a 0.7% agarose gel and DNA visualised by ethidium bromide staining.

2.6 Protein techniques

2.6.1 Polyacrylamide gel electrophoresis (PAGE) of protein samples

Polyacrylamide gels were used for the separation and analysis of proteins. All gels had an acrylamide to bis-acrylamide ratio of 29.5:1 (Boehringer Mannheim). Gels consisted of a 5% acrylamide stacking gel portion made up in 125 mM Tris, pH 6.8, 0.1% SDS, 0.05% AMPS and 0.001% TEMED. The resolving part of the gel, 12.5% acrylamide, was made up in 374 mM Tris, pH 8.8, 0.1% SDS, 0.1% AMPS and 0.001% TEMED. Samples were diluted 1 in 2 with protein sample buffer (62.5 mM Tris, pH 6.8, 2.5% glycerol, 25 mM dithiothreitol [DTT], 0.025% bromophenol blue), and heated at 95°C for 2 minutes before loading onto the gel. Gels were run in 25 mM Tris, 259 mM glycine, 0.1% SDS at 20 mA until the bromophenol blue dye front reached the bottom. Sea Blue molecular weight size markers (Norex), were also run on the gel (See Appendix 8 for molecular weight sizes).

Proteins which had been separated by SDS-PAGE were stained in the gels with Coomasie Brilliant Blue R250. The gels were submerged in staining solution (45% methanol, 10% glacial acetic acid, 0.25% Coomasie Brilliant Blue R250) and left to rotate gently on a platform shaker for 4 hours at room temperature. The gels were removed and placed in destaining solution (45% methanol, 10% glacial acetic acid) with gentle shaking for 4-18 hours. The gels were photographed and/or stored in 10% acetic acid.

2.6.2 Metabolic labelling of proteins produced by the recombinant baculoviral vector

Sf9 (0.5×10^6) cells were seeded onto 7 wells of a 24-well cell culture dish (COSTAR). Cells were infected with 10 pfu/cell of the recombinant virus for 1 hour. The inoculum was replaced with 1 ml of complete medium and returned to 27°C for 3 days. The complete medium was subsequently removed from infected cells and replaced with 0.5 ml of methionine- and cysteine-free TC 100 medium (GIBCO BRL), and returned to 27°C for 1 hour. The medium was aspirated and replaced with 150 µl of methionine- and cysteine-free TC 100 medium containing 10 µCi of Trans ^{35}S-methionine and -cysteine labelling mix (Amersham; 1000 µCi/mmol; 1 Ci=37GBq), at room temperature for 2 hours. After labelling, the radioactive medium was removed and cells solubilised in 100 µl lysis buffer (62.5 mM Tris-HCl, pH 6.8; 2.5% SDS; 12.5% glycerol; 0.1 M DTT and 0.005% bromophenol blue). Cell lysates were heated at 95°C for 2 minutes and then electrophoresed on a 12.5% SDS-PAGE (see section 2.6.1). Gels were transferred onto 3 MM paper and dried under vacuum at 70°C for an hour. Proteins were visualised by autoradiography and molecular weights were estimated by comparing electrophoretic mobility with that of Sea Blue molecular weight size markers (See Appendix 8 for molecular weight sizes).

2.6.3 Purification of Cre recombinase from immobilised Ni^{++}-spin column

A metal chelate adsorbent, Ni^{++}-NTA (nitrilo-tri-acetic acid, charged with nickel ions) with a strong binding affinity ($K_d=10^{-13}$, pH 8.0; far greater than the affinity between most antibodies and antigens, or enzymes and substrates) for proteins or peptides containing six or more consecutive histidine residues at either their N- or C-terminus (Hochuli *et al.*, 1988), was used for the purification of proteins containing histidine tag in a single-step by affinity chromatography.

Sf9 (5×10^9) cells were infected with the recombinant baculovirus at 10 pfu/cell. Sf9 infected cells were harvested 3 days after the infection by centrifugation at 1000 rpm (~ 135 g) for 5 minutes and washed twice with PBS. The cell pellet was resuspended by vortexing in 5 ml of cold binding buffer (4 mM NaH_2PO_4; 46 mM Na_2HPO_4; 300 mM NaCl; 2.5 mM Imidazole, pH 8; to which 10 mM β-mercapthoethanol and 1 × protease inhibitor cocktail [Pancreas extract 15 µg/ml; Pronase 1.5 µg/ml; Chymotrypsin 1.5 µg/ml; Thermolysin 0.8 µg/ml; Trypsin 0.002 µg/ml and Papain 1 mg/ml] were added just before use). The cell suspension was freeze-thawed three times and centrifuged at 14000 rpm (10508 g) for 10 minutes. The supernatant was removed and loaded onto Ni^{++}-NTA spin column (prepared by adding 2 ml of Ni^{++}-NTA agarose [Qiagen] to a 10 ml column and then centrifuged at 1000 rpm [~135 g] to pack the column), which had been equilibrated with 5 ml of the binding buffer. The column was erected to allow protein(s) to bind under gravity for 10 minutes and then centrifuged at 1000 rpm (~135 g) for 15 minutes to remove unbound proteins.

The column was subsequently washed twice (to remove proteins that bound non-specifically to the NTA resin) with 4 ml of wash solution (4 mM NaH_2PO_4; 46 mM Na_2HPO_4; 30 mM NaCl; 5 mM Imidazole; pH 8, 10 mM β-mercapthoethanol and 1 × protease inhibitor cocktail) by centrifugation at 1000 rpm (~135 g) for 15 minutes between washes. After the second wash, 1.5 ml of elution buffer (0.4 mM NaH_2PO_4; 4.6 mM Na_2HPO_4; 30 mM NaCl; 50 mM Imidazole; pH 8, 10 mM β-mercapthoethanol and 1 × protease cocktail inhibitors) was added to the column. Hexa-histidine-tagged (Cre recombinase) protein was eluted from the column by centrifugation at 1000 rpm (~135 g) for 15 minutes. The eluted sample was dialysed against 2 litres of Cre reaction buffer (20 mM Tris-HCl; pH 7.5, 300 mM NaCl and 1 ml Na_2EDTA) overnight at 4°C. The dialysed sample was filtered through 0.22 µM filter (non-pyrogenic, COSTAR), and stored as 100µl aliquots at -70°C or in 50% glycerol at -20°C. Aliquots (20 µl) were taken at each step in the purification process and 30 µl of protein sample buffer (62.5 mM Tris-HCl, pH 6.8; 2.5% SDS; 12.5% Glycerol; 25 mM DTT and 0.025% bromophenol blue) added prior to electrophoresis. Proteins were separated on a 12.5% SDS-PAGE and protein bands detected by Coomasie blue staining (see section 2.6.1).

2.6.4 Determination of protein concentration

Protein concentrations were determined by a colorimetric method based on the one described by Bradford (1976) using a commercially available reagent (Bio-Rad Protein Assay dye reagent).Protein samples (100 µl) were mixed with 5 ml of the 1 × reagent in a test tube. This was left at room temperature for 5-15 minutes before reading the absorbance at 595 nm. A standard curve was constructed using bovine serum albumin (BSA). The concentration of a protein sample was estimated by interpolating from the curve.

2.7 Protein-DNA interactions in Band Shift Assays

2.7.1 Non-denaturing polyacrylamide gels

All gels had an acrylamide to bis-acrylamide ratio of 29.5:1 (Boehringer Mannhein). Band shift assays were loaded onto 5% (w/v) non-denaturing polyacrylamide continuos gel made up in 0.5 × TBE buffer (89 mM Tris/borate, 89 mM Boric acid, 2 mM EDTA, pH 7.8), 0.1% AMPS, and 0.001% TEMED. Gels were pre-electrophoresed for 1 hour in 0.5 × TBE buffer and then run at room temperature for 2-3 hours at 30 mA. After electrophoresis gels were transferred onto 3 MM paper and dried under vacuum at 70°C for 1 hour. Band shifts were visualised by autoradiography.

2.7.2 Band shift assays

Band shift (or gel retardation) assays were done essentially as described by Zimarino and Wu, (1987) using recombinant protein purified from the insect-baculovirus system and double stranded oligonucleotides comprising DNA sequences.

2.7.2.1 Synthesis of oligonucleotides

All oligonucleotides were synthesised and purified by Hugh Spence (WRL [now GWR&D]). Purified oligonucleotides were dissolved in TE buffer and precipitated

with ethanol. After ethanol precipitation, the oligonucleotides were redissolved in appropriate volume of TE buffer and stored at -20°C. The concentration of oligonucleotide was determined by measuring it's optical absorbance at 260 nm given that an A_{260} of 1=20µg/ml of single stranded DNA. Alternatively the optical absorbance reading at 260 nm was used to calculate the molarity of the oligonucleotide using molar extinction coefficients of its components (extinction coefficients: T=9.7/base, C=9.2/base, G=11.4/base and A=15.4/base).

2.7.2.1 Annealing of complementary oligonucleotides

Oligonucleotide sequences were designed as sense and anti-sense strands to correspond to loxP, frt and loxP/frt hybrids sites (hybrid contains either loxP spacer or frt spacer sequences, see Appendix 4 for sequences). The complementary sense and anti-sense strands (2µg of each) were annealed in a 20 µl annealing buffer (10 mM Tris pH 8.5 and 50 mM NaCl). The solution was heated at 85°C for 5 minutes and left to cool slowly to room temperature.

2.7.2.2 5' labelling of oligonucleotide with ^{32}P

The 5' termini of the annealed oligonucleotides (0.1 µg) were radiolabelled by T4 polynucleotide kinase (Promega) at 2 units/µg of DNA in a 20 µl reaction volume containing 10 µCi [γ-^{32}P] dCTP (Amersham; 3000 Ci/mmol), and 1 × polynucleotide kinase reaction buffer (Promega). The reactions were incubated at 37°C for 30 minutes followed by inactivation of the enzyme by incubation at 70°C for 10 minutes.

2.7.2.3 Removal of unincorporated γ-^{32}P dCTP

Unincorporated nucleotides were removed by Sephadex G-25 gel filtration chromatography. Columns were prepared in 2 ml disposable syringes. The tip of the syringe was plugged with sterile siliconised glass wool. The column was packed with 2 ml of Sephadex G-25 pre-swollen in TE buffer. The column was placed in a 15 ml Falcon tube and packed by centrifugation 1000 rpm (135 g) for 2 minutes in a bench-top centrifuge. The column was washed with 500 µl of TE buffer and centrifuged as above. The kinased oligonucleotide sample was applied to the column and spun at 1000 rpm for 2 minutes. The radiolabelled oligonucleotide was collected into a sterile

microfuge tube which was placed inside the 15 ml Falcon tube. An estimate of the efficiency of incorporation was achieved by measuring the counts per second in solution before and after purification from the column using a Geiger counter. The radiolabelled oligonucleotides were stored as 10 µl aliquots at -20°C.

2.7.2.4 DNA-protein interaction in band shift assay

The ^{32}P-labelled oligonucleotides were incubated with or without the recombinant Cre protein in a total reaction volume of 10 µl containing the assay buffer (50 mM Tris-HCl, pH 7.4; 10 mM $MgCl_2$; 30 mM NaCl; 1 mg/ml BSA; 1 µg/ml poly dI-dC; 20% Ficoll). The reactions were started by the addition of the Cre protein and incubated at room temperature for 20 minutes. In the control reactions the interaction of the recombinant Cre protein and radiolabelled oligonucleotides were competed with 100-fold molar excess of the corresponding unlabelled oligonucleotides.

After incubation, reactions were loaded onto a 5% (w/v) non-denaturing polyacrylamide gel (see section 2.7.1). The gels were run at room temperature for 2-3 hours at 240V (30 mA). After electrophoresis, gels were dried and signals detected by autoradiography.

2.7.3 "On rate-off rate" band shift competition assays

The "on-rate" of Cre-loxP binding interaction was assayed by incubating the ^{32}P-labelled loxP oligonucleotide with Cre protein in a total reaction volume of 80 µl in the binding buffer (described above) at 37°C. At various time points, 10 µl of the reaction mixture was removed and run on a 5% non-denaturing polyacrylamide gel (as above).

The "off-rate" of Cre-loxP binding interaction was assayed by adding 10-fold molar excess of non-radiolabelled loxP oligonucleotide to the mixture after 10 minutes of pre-incubation of the Cre protein with radiolabelled loxP oligonucleotides ("on-rate"). At various time points into the competition reaction, 10 µl of the mixture was removed and immediately loaded onto the same gel as for the "on-rate".

2.8 *In vitro* assay systems

2.8.1 *In vitro* Cre-mediated recombination assay

Cre-loxP site-specific recombination *in vitro* was done by incubating the substrate plasmid (containing loxP sites) with or without the Cre protein in a 30 µl total reaction volume containing assay buffer (50 mM Tris-HCl, pH 7.4; 10 mM $MgCl_2$; 30 mM NaCl; 1 mg/ml BSA at 37°C for 20 minutes (unless otherwise stated). The reaction was stopped by heating at 70°C for 10 minutes. After heat inactivation, the recombination event was analysed in three ways:

2.8.1.1 Restriction digestion and Southern hybridisation analysis

DNA in the recombination reaction was separated from the Cre protein by phenol/chloroform extraction (see section 2.2.3). The DNA was recovered by ethanol precipitation (see section 2.2.2) and resuspended in 20 µl of distilled water. DNA was then digested with the restriction endonucleases *Bam*HI and *Sca*I and separated on a 0.7% agarose gel. The DNA fragments were then transferred onto a nylon membrane (Hybond- N, Amersham) and hybridised (as described in section 2.2.20) with a radiolabelled DNA fragment corresponding to the β-lactamase ampicillin resistance (amp^r) gene. Hybridisation signals were quantified using a phosphoimager (Molecular Dynamics).

2.8.1.2 Restriction digestion and quantification by Fluorescence

DNA in the recombination reaction was extracted with phenol/chloroform, recovered by ethanol precipitation and resuspended in 20 µl of distilled water as described above. The DNA was then digested with the restriction endonuclease *Sca*I and separated on an agarose gel. The gel was soaked in a SYBR Green solution (The stock solution was diluted 1:10,000 in TE buffer; Molecular Probes, Inc). The staining was done in the dark (by covering the staining container with aluminium foil) with gentle agitation at room temperature for between 10-40 minutes. The amount of DNA in each band was then quantified using fluorimager (Molecular Dynamics).

2.8.1.3 Analysis of recombination event by phenotypic blue/white colour selection

An aliquot of the recombination reaction (containing ~50 ng of DNA) was used to transform *E. coli* DH5α competent cells (see section 2.2.13) and dilutions plated onto L-agar (containing 100 µg/ml ampicillin 100 µg/ml IPTG and 150 µg/ml X-gal [5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside]) and incubated at 37°C overnight.

2.8.1.3b Analysing recombinants by sequencing

Single white colonies isolated on L-agar plates (see section 2.8.1.3) were picked into 10 ml of LB medium and grown overnight with shaking at 37°C. Plasmid DNA was prepared from these cultures using Promega Magic/Wizard mini-prep kit following the manufacturer's protocol. Plasmid DNA preparations were sequenced as described in section 2.2.21.

2.8.2 CAT ELISA assay of transformed CHO-DUKX B11 cells

Enzyme-linked immunosorbant assay (ELISA) was used for the quantitative determination of chloramphenicol acetyl transferase (CAT) in transfected CHO-DUKX B11 cells. The CAT ELISA was done following the manufacturer's protocol (Boehringer Mannheim).

The CAT ELISA is based on the sandwich-immunoassay principle. Antibodies to CAT (anti-CAT) were prebound to the surface of microtitre plate modules (MTP modules). Following lysis of the transfected cells, the cell extracts, which contained CAT were added to the MTP wells. All CAT contained in the cell extracts binds specifically to the anti-CAT antibodies bound to the plate surface. Next a digoxigenin-labelled antibody to CAT (anti-CAT-DIG), was added which binds to CAT. In the next step, an antibody to digoxigenin-conjugated to peroxidase (anti-DIG-POD) was added which binds to the digoxigenin. In the final step, the peroxidase substrate was added. The peroxidase catalyses the cleavage of the substrate yielding a coloured reaction product. The absorbance of samples were determined using a microtitre plate (ELISA) reader at 405 nm with a reference wavelength at approximately 490 nm.

2.8.3 Secreted Alkaline Phosphatase (SAP) assay

SAP assay was used for the quantitative determination of transfected CHO-DUKX B11 cells expressing the reporter gene encoding SAP. A 200µl aliquot of the culture medium was transferred to a microcentrifuge and centrifuged at 14000 rpm (15800 g) for 2 minutes, to pellet cell debris. The clarified supernatant was transferred to another microcentrifuge tube and heated at 65°C for 30 minutes to inactivate non-placental alkaline phosphatases present in the serum (SAP is stable at 65°C). The supernatant was diluted 1 in 100 (or less if low levels of SAP expected), in diethanolamine (DEA) buffer (1 M diethanolamine [Sigma], 0.5 mM $MgCl_2$, pH 9.6). Substrate solution made by dissolving one 20 mg tablet of ρ-nitrophenyl phosphate (Sigma) in 10 ml of DEA buffer (to give 2 mg/ml was ρ-nitrophenyl phosphate in DEA buffer) was dispensed into a 96 well microtitre plate at 180 µl per well. The assay (done in duplicates) was started by adding 20 µl of diluted culture supernatant (Time 0). The initial absorbance at 405 nm (Time 0) was determined by a microtitre plate reader. Assay plates were incubated at 37°C for at least 30 minutes and the A_{405} measured again. The absorbance of more than 1 unit was disregarded (because absorbance reading of more than 1 did not fit a linear curve) and the change in absorbance used as a measure of SAP activity.

2.8.4 Determination of Specific Production Rates (SPR)

The levels of SAP expression in transfected CHO-DUKX B11 cell lines were compared by determining the specific production rates of SAP, expressed in arbitrary units/10^6 cells/day (where arbitrary unit is defined as the change in absorbance in 30 minutes).

A T75 flask was seeded with 1×10^6 cells in 5 ml of culture medium and incubated overnight at 37°C. The medium was aspirated and replaced with 5 ml of fresh medium. After 24 hours of adding the fresh medium (± 30 minutes), 0.5 ml of medium was removed for SAP determination. The adherent cells in the flask were detached by versene + trypsin treatment, and viable cell density determined. Following this, the dislodged cells were diluted with 5 ml of fresh medium. Cell suspension was transferred to a 15 ml Falcon centrifuge tube (Becton Dickinson), and pelleted by centrifugation in a Sorvall RT6000B centrifuge at 1000 rpm (~ 135 g) for 5 minutes. The cell pellet was resuspended in 5 ml of fresh medium, an aliquot was

diluted with 9/10 volume of trypan blue solution (0.4% in saline, ICN Flow), and viable cells (non-stained) counted using a haemocytometer.

Specific production rate (SPR) was calculated as;

$$\text{SPR (arbitrary units/}10^6\text{ cells/24 hours)} = \frac{\text{Arbitrary unit of SAP activity}}{\text{cell density (}10^6\text{/ml)}}$$

2.8.5 β-gal staining of transformed CHO-DUKX B11 clones

β-gal staining was used as an in situ method to detect transfected CHO-DUKX B11 clones expressing activated *lac*Z (β-galactosidase) reporter gene.

CHO-DUKX B11 transformed colonies (growing on 10-cm petri dishes) were washed once with 10 ml of PBS (GIBCO BRL) after the culture medium has been aspirated. Transfected colonies were fixed onto the petri dish by the addition of 10 ml of 2% paraformaldehyde/0.3% glutaraldehyde in PBS solution. Fixation was done at room temperature for 15 minutes after which colonies were washed twice with PBS solution. Staining was done by adding 5 ml of a chromophore solution (consisting of 0.1% X-Gal [5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside; Promega], 5 mM $K_4Fe(CN)_6.3H_2O$, 5 mM $K_3Fe(CN)_6$, and 2 mM $MgCl_2$ in PBS; Promega) diluted 1:50 with X-Gal mixer solution (Promega) and incubated at 37°C for between 30 minutes and 18 hours. β-galactosidase expressing clones were observed as blue colonies under a light microscope.

2.8.6 β-galactosidase enzyme assay of transformed CHO-DUKX B11

β-galactosidase enzyme assay was used for the quantitative determination of β-galactosidase expression in transfected CHO-DUKX B11 cells. The assay was done using the β-galactosidase enzyme assay kit according to the manufacturer's protocol (Promega).

CHO-DUKX B11 transfected cells (1×10^6 cells) were harvested by versene-trypsin treatment. Cell suspensions were transferred to microcentrifuge tubes and centrifuged at 14000 rpm for 1 minute to pellet the cells. Cells were washed twice by resuspension

into PBS solution and centrifuged at 14000 rpm for 2 minutes in between washes. Cells were then freeze-thawed (freeze in dry ice and thaw at 37°C) three times in 1 × lysis buffer (Promega), to lyse the cells, vortexing after each thaw cycle. The cell extract was centrifuged in a microfuge for 5 minutes at 14000 rpm (10508 g) and the supernatant transferred to another microfuge tube.

An equal volume of 2 × assay buffer (120 mM Na_2HPO_4, 80 mM NaH_2PO_4, 2 mM $MgCl_2$, 100 mM β-mercapthoethanol, 1.33 mg/ml o-nitrophenol-β-D-galactopyranoside [ONPG], Promega) was added to 100 µl of the supernatant and the mixture incubated at 37°C for between 30 minutes and 6 hours during which time the β-galactosidase hydrolysis the substrate ONPG to o-nitrophenol (which is yellow). The reaction was terminated by the addition of 400 µl of 1M Na_2CO_3, mixed and 200 µl transferred to a 96 well microtitre plate. The absorbance of the samples were determined using a microtitre plate reader at 420 nm.

CHAPTER 3

Expression and purification of bacteriophage P1 Cre recombinase in insect cells using a recombinant baculovirus vector

3.1 Introduction

As was mentioned in the introduction to this project, one of the main objectives was to investigate the kinetics of the Cre-loxP site-specific recombination reaction *in vitro* with the aim of identifying irreversible substrates that could be used to improve the efficiency of gene targeting in mammalian cells. Another objective was to directly introduce Cre protein into cells to mediate targeted integration. Both objectives require a source of Cre protein. Therefore, to achieve these aims, the gene encoding Cre recombinase was expressed under the transcriptional control of the very strong baculovirus viral promoter, Ppolh (which normally drives the expression of the polyhedrin gene, described below) in insect cells infected with a recombinant baculovirus in the hope of obtaining high levels of biologically active Cre protein.

Escherichia coli (*E. coli)* has been used extensively as the host for the expression of a variety of prokaryotic and eukaryotic genes to produce large quantities of recombinant proteins (Pillot *et al.*, 1996; Amrein *et al.*, 1995; Itakura *et al.*, 1977); however, because of the highly reducing intracellular environment, over-expressed foreign proteins often fail to assume the proper tertiary structure or to be properly modified after translation. Consequently, the yields of biologically active proteins produced tends to be low. The baculovirus/insect expression system on the other hand, has the potential for the high expression of biologically active recombinant proteins because it provides a "eukaryotic environment" which is more suitable for gene expression including the necessary cellular post-translational modifications such as glycosylation, cleavage of signal peptide, disulfide-bridge formation, and proper folding.

As the work described in this chapter exploits baculovirus as an expression vector for Cre expression in insect cells, the biology of baculoviruses and the use of baculoviruses as expression vectors for heterologous gene expression in insect cells are described.

3.2 The biology of baculoviruses

Baculoviruses are a diverse group of viruses found mostly in insects. They are not known to have any non-arthropod hosts (reviewed in Miller, 1988a; Griffiths, 1994). The baculo portion of the name refers to the rod-shaped capsids of the virus particles. The following section describes the genetic organisation of baculoviruses.

3.2.1 The genetic organisation of baculoviruses

The DNA genome of a baculovirus is double-stranded, covalently closed and circular (Summers and Anderson, 1972). The DNA of the two baculoviruses commonly used for expression vector work, *Autographa californica* nuclear polyhedrosis virus (AcNPV) and *Bombyx mori* nuclear polyhedrosis virus (BmNPV) are both approximately 130 kbp. The baculovirus genome is compact in nature, this is exemplified by the observation that the translational termination signal (UAA) of an open reading frame (ORF) often overlaps the AAUAAA polyadenylation signal of the transcript. Similarly, promoter elements of ORFs may be found in the coding sequence of an upstream ORF (e.g., up39 and cg30 or p26 and p10). Furthermore, the sequences between many ORFs are extremely A + T rich, which may be related to promoter and transcriptional termination functions. The available information concerning the biological activity of some characterised essential genes for viral replication are summarised in the table below.

Class	Designation	Function
VL	*polh*	Polyhedrin, required for virus coat formation
L	*p74*	essential for occluded virus (OV) infectivity
L or VL	*fp25*	involved in nuclear envelopment process of OV
L or VL	*clx*	involved in the stability of polyhedrin inclusion bodies (PIBs)
L	*vp39*	major capsid protein

E and L (BV)	*gp*64	major envelope glycoprotein of budded virus involved in virus infection
E DNA	*dnapol*	DNA polymerase homologue, involved in viral replication
E	*pcna*	stimulates DNA replication and late gene expression
E	*hel*	helicase homologue, required for DNA replication
E	*p*35	blocks apoptosis
L	*ubi*	ubiquitin homologue, may be involved in the stability of heterologous gene products.

E, L, and VL indicates that a gene is transcribed in the early, late, or very late phases of infection respectively.

Some of the other genes encoded by baculoviruses appear to be non-essential for replication as determined by the ability of viruses with null mutations in the given gene to replicate in both cell culture and insects. Genes that have been characterised to be non-essential include; *egt* (O'Reilly and Miller, 1989), the *da*26 ORF (O'Reilly *et al.*, 1990) and *p*10 (encoding the p10 protein that constitutes the fibrillar structures [Kuzio *et al.*, 1984]). However, these non-essential genes may have other functions, for instance a role for p10 in the melting or disintegration of the larval host has been proposed and also deletion of p10 is reported to prevent insect cell lysis at very late times post infection (p.i), (Williams *et al.*, 1989). Several genes that appear to be involved in gene regulation have also been identified. These genes include those encoding IE-0, IE-1, IE-N, PE-38 and CG 30. IE-0 and IE-1 are able to transactivate some early promoters in transient assays (Guarino and Summers, 1986; 1987). The other three proteins IE-N, PE-38 and CG 30 all share common unique structural motifs; an unusual, double zinc-finger like motif and a c-terminal leucine zipper (Thiem and Miller, 1989b). Thus baculoviruses appear to have a family of related genes that are probably involved in gene regulation described in the next section.

3.2.2 Mechanism of gene regulation

Baculoviruses are transcribed in a regulated cascade. During wild-type NPV infection, at least three phases of viral gene expression can be distinguished; early (0-6 hours post infection [h. p. i.]), late (6-20 h. p. i.) and very late (20-72 h. p. i.). Most viral genes are transcribed primarily during one phase although some genes are transcribed in two or possibly three phases. Early genes may be defined as those genes that are transcribed in the absence of any other viral gene expression. Late and very late genes are dependent on early viral gene expression and on DNA replication. The late and very late promoters are distinctive and unusual with regard to their location. The primary determinant of both late and very late promoter activity is the tetranucleotide TAAG, which is located at the transcriptional start site of all known late and very late transcripts (Thiem and Miller, 1989a).

The different promoters are distinguished by their different activities during the three phases of infection. For instance, the *polh* promoter is an example of a very late promoter; it shows low activity during the late phase (i.e., 6 h. p. i. - 18 h. p. i) but becomes highly active beginning about 18 h. p. i. By 27 h. p. i. to 48 h. p. i., approximately 20% of the total polyadenylation RNA in the cell is *polh* mRNA (Adang and Miller, 1982). In contrast, the late *vp*39 promoter is most active between 12 h. p. i. and 24 h. p. i. (Thiem and Miller, 1990).

The biology of this infection process is what underlies the utility of baculoviruses as expression vectors (Miller, 1981) for heterologous gene expression in insect cells, as such a brief overview of this infection process in insect cells is given in the following section.

3.3 Process of infection in insects

The natural cycle of infection by baculoviruses in insect larvae is summarised in Figure 3.1. Two distinct forms of viruses are involved in the infection; occluded virus (OV) enveloped in a polyhedrin protein coat to a form known as polyhedrin inclusion bodies (PIBs) and the budded viral (BV) forms. PIBs provides protection of the embedded virus particles during the horizontal transfer among larvae and a means of

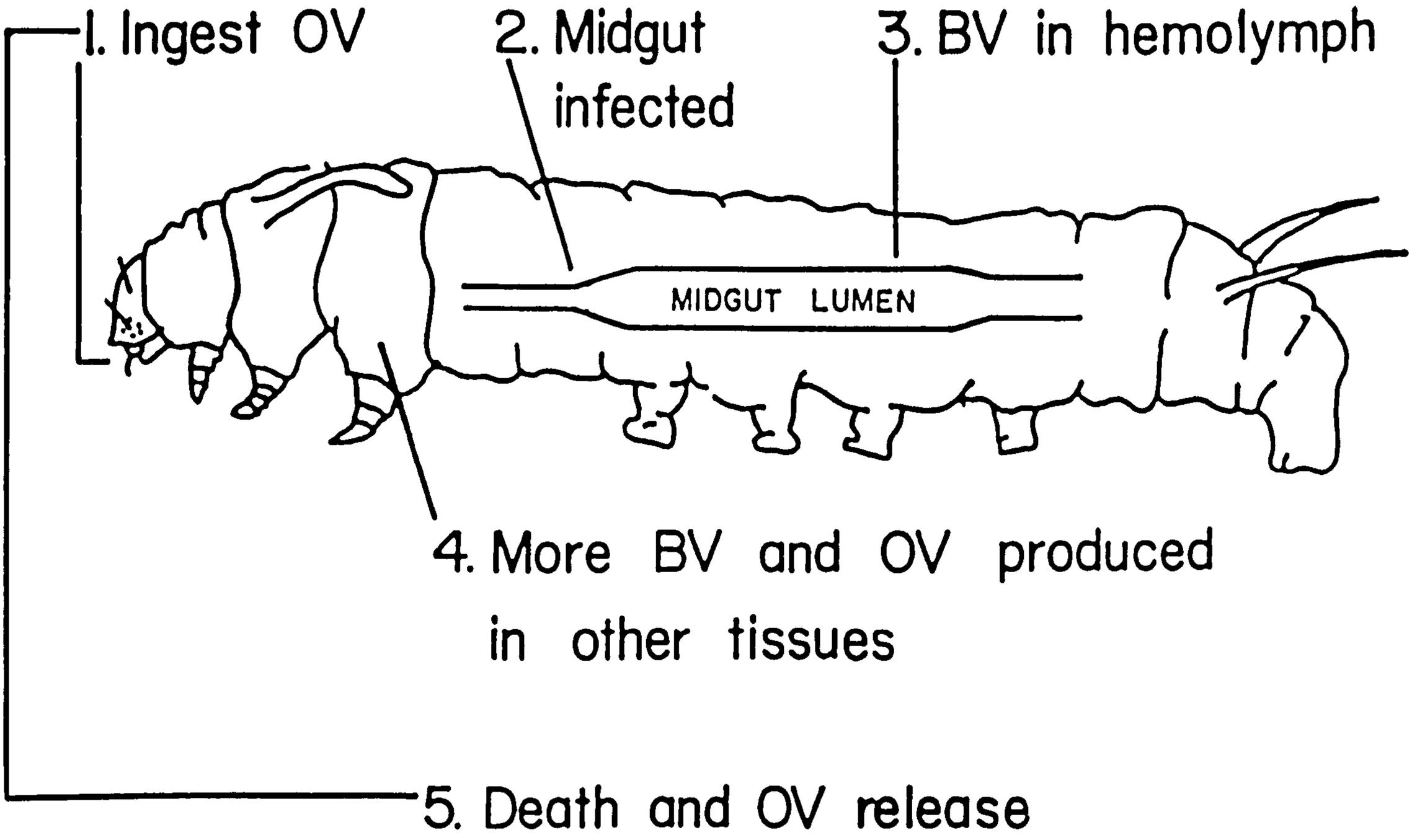


Figure 3.1- The natural cycle of infection by baculoviruses in insect larvae.
Two different forms of baculoviruses are involved in the infection process; the occluded viral (OV; also known as occlusion bodies) and budded viral (BV) forms. Occlusion bodies provide: (1) protection of the embedded virus particles during the horizontal transfer among larvae and (2) a means of delivering the virus particles to the primary site of infection, the midgut epithelial cells. The alkaline midgut facilitates dissolution in the midgut lumen thereby releasing the virus particles. BV released into the haemolymph cause the spread of infection to other tissues of the insect. Most larval tissues are susceptible to infection, and eventually the insect dies, the cuticle is disrupted, copious amounts of OV are released onto the plant surface, and another insect encounters the OV as a contaminant of its food.

delivering the virus particle to the primary site of infection, the midgut epithelial cells. BV which is released from the primary site of infection causes the spread of infection to other tissues of the insect.

Insect larvae ingest PIBs as contaminants of their food. The crystalline polyhedrin coat is solubilised in the alkaline midgut of the insects, releasing embedded virions (Harrap and Longworth, 1974). The virions enter midgut cells by fusion with the membrane of the microvilli (Granados and Williams, 1986a). Infection in the polarised midgut cells results in BV release from the basement membrane side of the cell causing the spread of infection to other tissues of the insect. This BV can gain access to the hemocoel and is transported via the hemolymph to other tissues in the insect (insects have an open circulatory system). BV released from the midgut also infects the epithelial cells of tracheoles, which provide oxygen to the midgut, spreading the infection along the tracheal network (Keddie *et al.*, 1989). During a typical BV infection, the insect continues to feed during most of the infection process, which takes approximately five to seven days. The integument becomes swollen and changes in lustre. Both BV and PIBs are produced in most of the tissues infected during the secondary phase. Infection of a late instar larva probably involves approximately 10 generations of virus. Eventually, the insect becomes lethargic and stops feeding. The cuticle melanizes due to a polyphenol oxidase-mediated process that results in a discoloration (browning). The musculature disintegrates, and the larva becomes a cuticular sac of milky fluid containing PIBs. A substantial portion (e.g., ca. 25%) of the dry weight of the liquefied carcass is polyhedra. The cuticle eventually ruptures releasing the PIBs into the environment. Polyhedra are relatively stable in the environment, although they exhibit significant sensitivity to UV light. They are naturally dispersed by a variety of routes and may eventually be consumed by another permissive insect host, thereby reinitiating the infection cycle (Evans, 1986).

The polyhedrin coat, whilst being important for the horizontal transfer of virus to new insect hosts, is not important for the propagation of virus in insect cell culture. Hence the polyhedrin (and p10 proteins), which comprise up to 50% total cell protein synthesised at late stage of infection, are not essential for virus replication *in vitro*. Understanding the biology and the molecular basis of the baculovirus infection process has allowed the manipulation of the baculovirus genome for the expression of heterologous proteins in insect cells (Luckow and Summers, 1988; Smith *et al.*, 1983a) which is described in the next section.

3.4 Use of baculovirus infected insect cells for the expression of heterologous proteins

As previously mentioned, two of the highly expressed proteins in insect cells infected by baculoviruses, namely polyhedrin and p10 are not essential for the propagation of the virus in insect cell culture. Therefore, recombinant baculovirus expression vectors can be constructed by the replacement of one of these non-essential genes, usually polyhedrin gene (*polh*), with the coding region of the heterologous protein so that high levels of expression can be achieved under the transcriptional control of the strong polyhedrin promoter following infection.

3.4.1 Development of recombinant baculovirus expression vectors

The basic design of the first baculovirus gene expression vectors involved the replacement of *polh* with the heterologous gene under *polh* promoter control (Miller *et al.*, 1983b; Smith *et al.*, 1983b; Pennock *et al.*, 1984; Maeda *et al.*, 1985). One advantage of substituting the heterologous gene in place of *polh* is that such recombinant viruses replicate normally as BV in cell culture and can be distinguished visually from wild-type (wt) virus: as they lack the crystalline coat polyhedrin protein, they appear round and clear (white) compared to the wild-type virus which appears as opaque clusters of crystals on an agarose overlay assay under a light microscope. Such recombinant viruses are not efficient at infecting larvae by the natural oral route of infection and do not persist in the environment; these features have some benefits from environmental and recombinant DNA safety perspectives (Miller, 1981).

The two baculoviral DNAs commonly used as expression vectors are *Autographa californica* nuclear polyhedrosis virus (AcNPV) and *Bombyx mori* nuclear polyhedrosis virus (BmNPV). AcNPV genome (~130 kbp) is too large for the efficient introduction of the foreign gene by conventional DNA cloning techniques. Consequently a two step procedure is followed in which the gene to be expressed is first inserted into a transfer vector containing a viral expression cassette flanked by viral DNA sequences which are required for the targeting of the gene to the viral genome. This recombinant transfer vector is co-transfected with viral genomic DNA into host insect cells. Subsequently host enzymes in the cells catalyse homologous recombination between the recombinant transfer vector and the viral DNA which

leads to the targeted integration of the foreign gene into the polyhedrin locus. These recombinant viruses can be selected or screened for described in section 3.4.1.2.

3.4.1.1 Transfer vectors

Transfer vectors (reviewed in O'Reilly *et al.*, 1992 and Bishop, 1992) consist of a fragment of AcNPV DNA, usually the *Eco*RI "I" fragment which contains the polyhedrin gene promoter, coding region and polyadenylation sequence flanked by other viral sequences, cloned into a pUC based plasmid vector. A multiple cloning site containing unique restriction endonuclease sites is usually inserted downstream of the promoter for the cloning of heterologous genes.

Early transfer vectors directed the synthesis of a fusion protein, containing the amino terminus of the polyhedrin protein, by inserting the heterologous gene downstream and in frame of the polyhedrin transcriptional initiation codon (ATG). Later vectors have the polyhedrin transcriptional initiation codon mutated to ATT or ATC (e.g. pVL941; Luckow and Summers, 1989; p36c; Page, 1989), so that transcription starts at the ATG codon of the inserted cDNA to produce an authentic, nonfusion protein. Other improved derivatives of the initial transfer vector has been described elsewhere (Matsuura *et al.*, 1987; Peakman *et al.*, 1992a; Belyaev and Roy, 1993).

3.4.1.2 Advances in the isolation of recombinant baculovirus

Originally, recombinant baculoviruses were identified phenotypically as polyhedra negative because the polyhedrin protein is not expressed, with lytic plaques appearing clear rather than opaque on an agarose overlay assay. However, screening was made difficult by the fact that the recombination event is relatively inefficient with recombinant polyhedrin negative clones only making up 0.1-1% of progeny virus. Various strategies have been developed to increase the percentage of recombinants among the progeny viruses and to aid their identification (reviewed in Davies, 1994). Several modifications of this procedure have been described that facilitate the identification of recombinant viruses by placing a reporter cassette adjacent to the gene to be expressed (Vialard *et al.*, 1990; Vlak *et al.*, 1990) or that increase the proportion of the viruses that are recombinants (Peakman *et al.*, 1989; Kitts *et al.*, 1990).

Kitts *et al.*, (1990) engineered an AcNPV derivative, AcRP6-SC, that contained a single *Bsu*36I restriction site downstream of the polyhedrin transcription start site. They showed that linearised viral DNA has a greatly reduced infectivity (one fifteenth) in comparison to the circular form but is still able to recombine with a co-transfected transfer vector to yield circular recombinant virus. Hence the use of linearlised virus reduces the background of non-recombinant parental viruses, resulting in (30%) of progeny viruses being recombinant (rather than 0.1-1%). The remaining non-recombinants result from incomplete *Bsu*36I digestion of AcRP6-SC or self-circularisation of linearlised AcRP6-SC.

A further improved selection system, resulting in ~95% of progeny virus being recombinant, was described by Kitts and Possee (1993). They engineered an AcNPV derivative, BacPAK6 (Figure 3.2), which contains three *Bsu*36I restriction sites near the polyhedrin locus. One is in the *lac*Z gene under the control of the polyhedrin promoter, another is in a non-essential gene (ORF 603) upstream of the polyhedrin promoter and the third is in the essential gene (ORF 1629) downstream of the polyhedrin gene. When *Bsu*36I digested BacPAK6 is used as the parental virus, only a double recombination event between the large *Bsu*36I fragment and a transfer vector will restore the essential ORF 1629 gene and create a circular virus capable of replication. This results in the majority (~95%) of the viable progeny produced being recombinant. Furthermore, double recombinants lack the *lac*Z gene, hence viral progeny (observed as plaques) will appear colourless rather than blue after incubation with medium containing X-gal. The frequency of non-recombinants is reduced to approaching 0% by the more efficient linearlisation due to the presence of three *Bsu*36I sites and the fact that self-ligation of three virus fragments to a form that can replicate is unlikely.

More recently, use of herpes virus thymidine kinase (*tk*) enzymes has allowed drug selection of recombinant viruses. An AcNPV-derivative has been engineered that contains the herpes simplex virus type 1 *tk* gene under the control of the baculoviral immediate early (IE-1) promoter, active in the immediate early phase of infection (Godeau *et al.*, 1992), making the virus sensitive to the anti-herpes drug, Gancyclovir. A recombination event between the virus and a transfer vector results in progeny viruses that lack the *tk* gene and can be selected over parental virus by passaging in the presence of Gancyclovir.

Alternative, *in vitro* methods for the creation of recombinant viruses have been described, for example, the use of the bacteriophage P1 Cre-loxP system. A loxP site

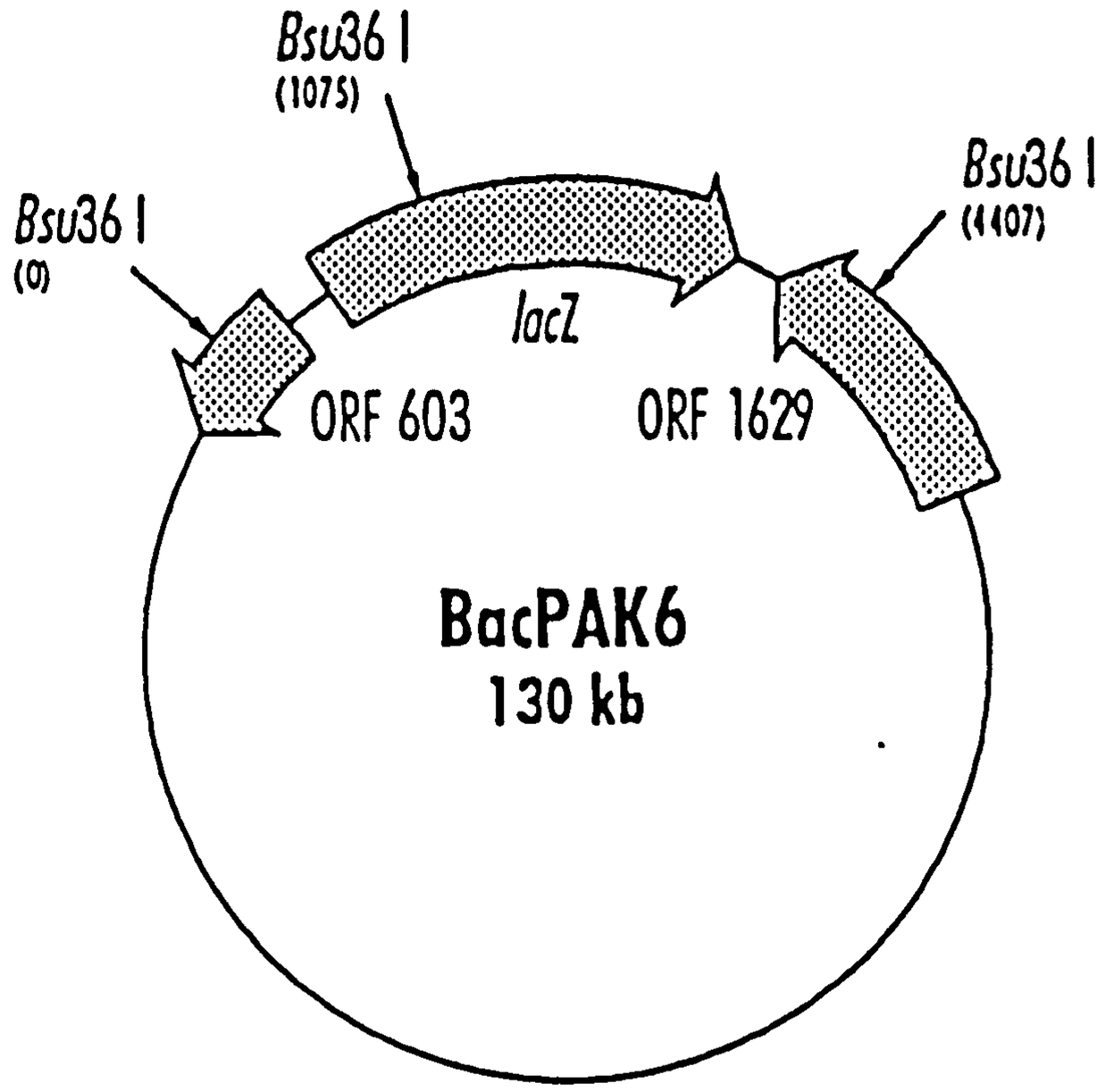


Figure 3.2-Simplified restriction map of BacPAK6 (Clontech, labs. Inc).

BacPAK6, a derivative of AcMNPV, is engineered to contain three *Bsu*36I sites; one in the *lacZ* gene under the control of the polyhedrin promoter, another in a non-essential gene (ORF603) upstream of the polyhedrin gene, and the third in the essential gene (ORF1629) downstream of the polyhedrin gene. Digestion of BacPAK6 with *Bsu*36I releases two small fragments, one of which contains part of the ORF1629 gene which is essential for viral replication. If the large fragment recircularises by itself, part of ORF1629 is missing and so the viral DNA cannot give rise to viable viruses. If however it recombines with a transfer vector containing the missing sequences, the resulting circular viral DNA can produce viable viruses.

inserted into the viral genome and the transfer vector allows recombination at high frequency in the presence of Cre recombinase (Peakman *et al.*, 1992a). However, although the probability of successive recombination events, leading to unwanted multiple insertions and rearrangements is high, appropriate recombinant viruses are easily identified by simple polymerase chain reaction (PCR) protocols. Furthermore, this is the only system that is independent of the size of the inserts.

3.4.2 Recombinant protein expression

As was mentioned previously, the baculovirus/insect cell system is an attractive alternative to *E. coli* as the host for foreign protein expression because protein expression levels can be very high, with some heterologous proteins comprising up to 50% of total cell protein (Miyamoto *et al.*, 1985).

Furthermore insect cells expresses enzymes required for post-translational modifications. Post-translational modifications observed in the system include glycosylation (up to high mannose), fatty acid acylation, amino-terminal acetylation, carboxy-terminal α-amidation and phosphorylation (although differences in the extent of modifications in insect and mammalian cells have been observed, e.g. glycosylation and phosphorylation [Hoss *et al.*, 1990]). In most cases, a recombinant intracellular protein is targeted to its appropriate intracellular location in the insect cell, for example to the nucleus (human c-myc; Miyamoto *et al.*, 1985). Insect cells are capable of recognising and cleaving both insect and mammalian signal peptides that direct proteins to the ER (Jarvis *et al.*, 1993), after which correct targeting to the cell surface (influenza virus haemagglutinin; Possee, 1986) or secretion (human beta interferon; Smith *et al.*, 1983) are observed. Other specific proteolytic cleavages have been observed, for example activation of influenza virus haemagglutinin by proteolytic cleavage of the precursor into two fragments (Possee, 1986).

For the expression of a protein consisting of more than one polypeptide chain insect cells may be co-infected with multiple recombinant viruses, each containing a single foreign gene, or a virus containing multiple foreign gene expression cassettes may be constructed. Early vectors expressing two foreign genes contained two copies of the polyhedrin promoter in opposite orientation. Vector instability due to homologous recombination between the two copies led to the development of vectors containing different promoters e.g. the polyhedrin and p10 promoters (reviewed in Griffiths, 1994). Both hetero- and homo-oligomeric assembly have been demonstrated. For

example, Belyaev and Roy (1993) co-expressed four bluetongue virus proteins in insect cells using a quadruple recombinant virus, which assembled into bluetongue virus-like particles.

3.4.3 Limitations to the baculovirus/insect cell expression system

Although the baculovirus/insect expression system is increasingly being used for the expression of heterologous proteins, a number of limitations of this system has been identified which are described in this section.

3.4.3.1 Expression levels

Most heterologous proteins are not expressed at levels comparable with those of the polyhedrin protein in a wild-type infection. In addition, expression levels of different proteins vary widely, with recombinant proteins having been produced as fusion or non-fusion proteins at levels ranging from 1 to >500 µg/ml. Expression levels also appear to be protein-specific, because different genes inserted into the same vector are expressed at different levels and the stage at which protein expression is limited seems to vary depending on the protein (Luckow and Summers, 1988; O'Reilly *et al.*, 1992).

Hasnain *et al.*, (1994) observed a difference at the level of transcription and suggested that the secondary structure of the foreign gene may influence the efficiency with which it is expressed. They expressed two proteins; firefly luciferase (luc) and the β-subunit of human chorionic gonadotrophin (βhCG), using a double recombinant baculovirus, both under the control of the polyhedrin promoter and found that luc mRNA (and synthesised protein) levels were higher than βhCG mRNA levels. Sequence analysis indicated extensive secondary structure and stem-loop complex-forming potential for the βhCG gene, which could be responsible for the transcriptional difference observed.

In some cases, the limiting step appears to be post-transcriptional. For example, Carbonell *et al.*, (1988), found that the levels of scorpion toxin mRNA (a fusion containing the first 58 codons of the polyhedrin gene) were equivalent to polyhedrin mRNA in a wild-type infection, but foreign protein levels were much lower than the polyhedrin protein level. In this case it appears that recombinant protein expression is

limited by mRNA stability and/or translation efficiency and/or protein stability. However, the recombinant toxin protein produced could also be toxic to the cell.

Paradoxically, although a codon bias exists in the highly expressed late baculovirus proteins, this may not be a major limitation to the translation of foreign mRNA because *E. coli* β-galactosidase (*lac*Z) gene which contains many codons under-represented in the late proteins, is still expressed at a very high level (200 μg/ml, Jarvis *et al.*, 1990).

Low expression levels may be attributable to the pathways through which the protein is co-and post-translationally processed in the cell; secreted glycoproteins are particularly poorly expressed compared to cytoplasmic or nuclear proteins. It was initially thought that heterologous signal peptides might be inefficiently recognised by the protein translocation machinery in *lepidopteran* insect cells, and evidence for this was presented by Tessier *et al.*, (1991), who showed that secretion of plant propapain was enhanced by replacing its native signal peptide with that from honeybee prepromellitin. However, this is not the case for all proteins, for example Jarvis *et al.*, (1993) found that low human tPA secretion (<5μg/ml) was not enhanced by an insect-cell derived signal peptide because similar yields were obtained from the insect secretory protein (tPA) containing a baculovirus glycoprotein-signal peptide. They concluded that some other factor, downstream of protein translocation into the ER, was limiting tPA secretion. This limitation could be glycosylation and/or protein folding and/or disulphide bond formation particularly if the cell's (organelle) membranes are highly compromised due to late stage lysis.

3.4.3.2 Quality of recombinant protein

The polyhedrin promoter is generally utilised in the baculovirus/insect cell expression system due to its very high activity (polyhedrin protein making up to ~50% of the total cell protein in wild-type baculovirus-infected insect cells). In addition, its activity late in viral infection, after the virus DNA has replicated (replication begins at ~6 hours p.i.), means that foreign protein synthesis should not interfere with virus replication and the increased foreign gene copy number within the cell contributes to the high protein expression level. The other very late promoter, p10, has also been used to drive foreign gene expression, and has similar activity as the polyhedrin promoter.

However, for some proteins, transcription initiated by one of the very late, strong promoters, results in only a small proportion of synthesised protein being biologically active and/or correctly processed. For example, Aris *et al.*, (1993) found that the intracellular protein, protein kinase C-δ was expressed at a relatively high level in insect cells, comprising 10-20% of total cell protein, but that less than 1% of the material was catalytically active, the majority aggregating in insoluble complexes. In a different study, Jarvis and Summers (1989) reported that whilst secreted human tPA produced in a recombinant baculovirus/insect cell system was glycosylated, the processing of terminal residues was incomplete, with some of the material having a complex carbohydrate addition and the remainder having a high-mannose pattern.

These and numerous other reports of incomplete protein processing suggest that when expression is controlled by a very late promoter, leading to relatively high levels of the foreign protein (up to 50% total cell protein, in some cases), the insect cell's post-translational modification machinery becomes saturated plus cell membranes are compromised. Furthermore, there may be insufficient time for the complete processing of the protein from its synthesis, very late in infection, to insect cell lysis. This situation may be exacerbated by the fact that the majority of insect cell-specific protein synthesis is switched off in the very late stage of viral infection, possibly resulting in a compromised processing and secretory pathway. A variety of approaches have been taken to overcome these limitations at the transcriptional regulation level, which are described in the next section.

3.4.3.3 Overcoming limitations by the use of earlier viral promoters

The use of promoters active at an earlier stage of infection has been investigated in an attempt to overcome some of the limitations associated with the use of the very late polyhedrin and p10 promoters, i.e. the incomplete processing of recombinant protein and low expression of secretory protein. Many early and late viral genes encode for essential proteins required for replication of the virus and so transfer vectors containing a copy of such an early or late promoter in the polyhedrin locus are used.

Late promoters are transcriptionally weaker than the very late promoters, but show peak activity at earlier times in infection (12-24 h. p. i.), when the cell's protein synthesis machinery is efficient for post-translational modification and secretion. In addition, proteins expressed from a late promoter will have an additional 12 hours to move through the Endoplasmic reticulum (ER) and the Golgi apparatus, compared to

expression from a very late promoter. One late phase promoter used for foreign gene expression is that of the basic protein, a core-protein tightly associated with the viral DNA (Hill-Perkins and Possee, 1990), which is expressed a high level of β-galactosidase, although lower than with the polyhedrin promoter. Rankl *et al.*, (1994) found that expression levels of protein kinase C-δ were four-fold lower from the basic protein promoter, compared to the polyhedrin promoter, but that the enzyme was 10 to 15 fold more active with the basic protein promoter. Thus the earlier expression of the protein resulted in the production of more active enzyme. It is not always the case that a late promoter results in a lower protein expression levels, compared to the very late p10 and polyhedrin promoters. Bonning *et al.*, (1994) showed that the expression levels of two enzymes, β-galactosidase and juvenile hormone esterase, were higher for the late phase basic protein promoter than the two very late promoters, p10 and polyhedrin, both in terms of amount of protein and activity of the protein. This is perhaps due to the different phases of maximal activities observed with the different promoters.

Whilst the use of early phase baculoviral promoters in recombinant baculoviruses to improve their insecticidal activity has been investigated (O'Reilly and Miller, 1989; O'Reilly *et al.*, 1990), their use to achieve higher levels of active recombinant protein has not been extensively studied. The immediate early 1 (IE-1) promoter is active immediately post-infection, with the IE-1 mRNA transcript being detected in wild-type virus-infected cells within 1 h. p. i. (Guarino and Summers, 1986). The IE-1 protein is implicated in the regulation of viral gene expression, including transactivation of other early genes. The IE-1 promoter was used by Sharp and Gewert (unpublished) for β-galactosidase expression, and active levels of the enzyme were found to be 1000 fold lower, compared to the use of the polyhedrin promoter. A similar study (Griffiths, unpublished) found that the difference in the active levels of a secreted protein, SAP, was considerably smaller, with only 10-20 fold lower levels for the IE-1 promoter compared to polyhedrin promoter. Whilst a higher level of active material for both these proteins was achieved with the very late polyhedrin promoter, the difference was greatly reduced for the secreted protein suggesting that for other highly processed secreted proteins, the IE-1 promoter may be equally, or even more efficient than the polyhedrin promoter. A potential limitation to the use of an early or late phase promoter is the risk of interference with virus replication.

Because proteins expressed under the transcriptional control of IE-1 promoter does not cause insect cell lysis, the IE-1 promoter could be used in the construction of a stable insect cell line for recombinant protein expression. Such a stable cell expression

system has been described (Jarvis *et al.*, 1990) however, the levels of protein expression was usually lower than the levels observed in a regular baculovirus expression system. With further development an insect-based cell expression system will probably compare well to mammalian expression system.

3.5 Results of Cre recombinase expression and purification in the baculovirus/insect system

As mentioned in the introduction to this chapter, the aim of this project was to use the baculovirus/insect expression system to express high levels of Cre protein for the investigation of the Cre-loxP reaction *in vitro* and also the direct introduction of Cre protein into mammalian cells to catalyse gene targeting.

In order to investigate the baculovirus/insect expression system for producing recombinant Cre protein, the *cre* gene was cloned into a transfer plasmid pAcH6N1 to generate pAcH6N1-Cre. Insect cells were co-transfected with pAcH6N1-Cre and BacPAK6 viral DNA digested with *Bsu*36 I. Recombinant viruses resulting from the co-transfection in insect cells were observed as colourless plaques after a neutral red staining in a standard plaque assay procedure and validated by PCR analysis using viral DNAs prepared from isolated plaques. Metabolic labelling was used to investigate recombinant virus mediated synthesis of Cre protein infected insect cells. Cre protein synthesised in insect cells infected with the recombinant viruses was purified from an affinity chromatography column. The purified protein was then assayed for Cre activity using a simple *in vitro* assay system. In addition, various parameters for the Cre protein reaction were studied and compared to native Cre recombinase of Bacteriophage P1. Finally the advantages of using baculovirus-insect system for producing Cre protein on a large scale for gene targeting and *in vitro* analysis are discussed.

3.5.1 Construction of chimeric transfer vector pAcH6N1-Cre

As previously noted, the baculovirus genome is too large and usually contains more than one recognition site for known restriction endonucleases. This makes it difficult to efficiently introduce the *cre* gene by conventional DNA cloning techniques. In order to generate a recombinant baculovirus encoding the Cre protein, the *cre* gene

Plasmid pAcH6N1 (Dirk Gewert, unpublished; Figure 3.3a) is derived from a AcMNPV-based transfer vector pVL1392 (Invitrogen). The pVL1392 transfer vector contains 6.5 Kb of the wild-type AcMNPV genomic DNA, including the polyhedrin gene promoter, Ppolh, protein coding sequence and polyadenylation signal sequences, and flanking viral DNA sequences. A polylinker for the insertion of foreign genes is located at the +35 with respect to the mRNA transcriptional initiation site at +1. The plasmid was designed in this way because several groups demonstrated that sequences downstream from the polyhedrin transcriptional initiation codon, ATG contributed to optimal promoter activity (Luckow and Summers, 1989; Page, 1989). The polyhedrin translational initiation codon, ATG was mutated to ATT, so that translation initiates at the ATG start site of the foreign gene. Recent evidence obtained with recombinant viruses generated with these plasmids indicated that despite the mutation at +3, some translation still initiates at the +1 position (Beames *et al.*, 1991).

Plasmid pAcH6N1 is derived from pVL1392 by replacing the polylinker sequence between *Bgl*II and *Kpn*I sites with oligonucleotide sequences encoding six histidine residues (the strong binding affinity of histidine residues for cations, means that heterologus genes fused to a hexa-histidine tag can be expressed and purified in a single-step by affinity chromatography; described in detail in section 3.5.4). The hexa-histidine sequence cloned into the *Bgl*II and *Kpn*I sites was out-of-frame with the ATT at +1, to ensure that translation would initiate at the first ATG of the histidine domain.

The 1.1 kbp *cre* gene encoded by bacteriophage P1 is engineered to have a nuclear localisation signal in plasmid pMC1-Cre (Gu *et al.*, 1993). By including the nuclear localisation signal, it was hope that this would increase the rate of transport of Cre protein into nucleus to improve the efficiency of site-specific recombination *in vivo*. The *cre* gene was amplified from pMC1-Cre in a PCR reaction. A combination of oligonucleotide sequences designed as forward (CB1741-gacaaG<u>AGATCT</u>**ATG**CCCAAGAAGAAGAGGAAGG) and a reverse (CB1742-gacaaG<u>AGATCT</u>**CTA**ATCGCCATCTTCCAGC) primers were used in the PCR reaction (Figure 3.3b). CB1741 anneals at the 5' end and CB1742 anneals at the 3' end of the *cre* coding sequence (see appendix 1). (The ATG initiation and the CTA [anti-sense] termination codons in CB1741 and CB1742 respectively are in bold. The introduced *Bgl*II sites are underlined and lower case letters indicates non-specific nucleotide

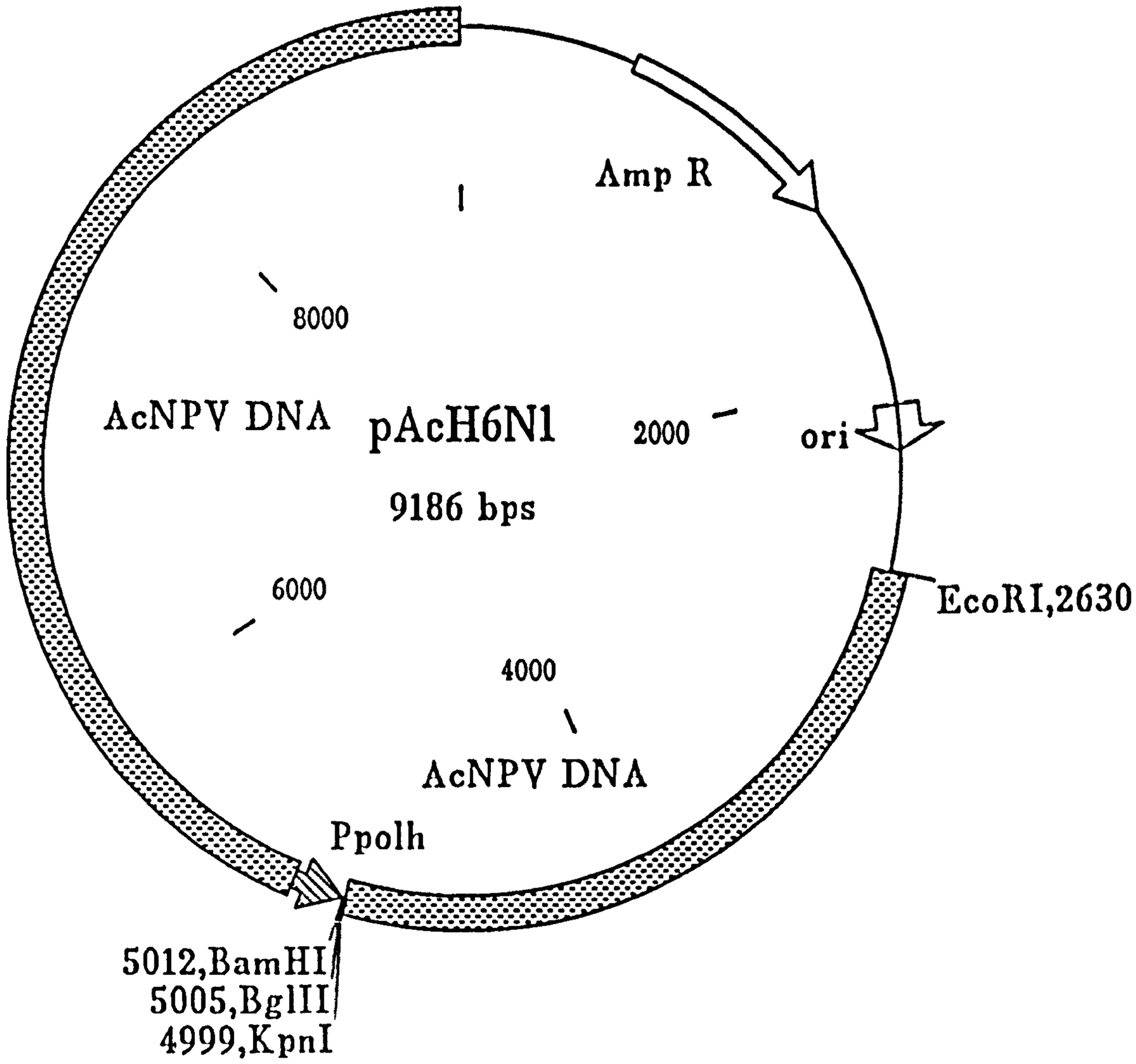


Figure 3.3a-Map of transfer vector pAcH6N1.

The transfer vector was derived from pVL1392 (Invitrogen). It contains 6 Kbp of the wild-type AcNPV genomic DNA (shown as dotted boxes), the polyhedrin promoter, Ppolh (shown as slashed box with arrow head) and the protein coding region including polyhedrin signal sequences (not shown). In addition it contains six histidine residues at position +35 with respect to mRNA transcriptional initiation site at position +1 (see Figure 3.3d).

extensions). The PCR product analysed by gel electrophoresis (Figure 3.3c) was purified and digested with *Bgl*II and subsequently cloned as a *Bgl*II fragment into the *Bam*HI site of plasmid pAcH6N1. The resulting plasmid was designated pAcH6N1-Cre (Figure 3.3d). Plasmid pAcH6N1-Cre, with the *cre* gene inserted in the correct orientation and frame for translation was validated by restriction endonuclease digestions and by sequencing.

The *cre* gene was cloned into pAcH6N1 downstream of the histidine residues, such that the translational initiation codon of the histidine domain would be in frame with that of the cre gene. The resulting transfer vector pAcH6N1-Cre encodes an mRNA species, containing an extended 5' untranslated sequence plus the first 33 bp (corresponding to the polyhedrin N-terminal coding sequence) in addition to the histidine and Cre fusion gene under the transcriptional control of the polyhedrin promoter.

3.5.2 Production of recombinant baculoviruses containing *cre* gene

In order to produce high proportion of viral progeny corresponding to recombinant baculoviruses containing the *cre* gene under the transcriptional control of the polyhedrin promoter, *Spodoptera frugiperda* (Sf9) insect cells (2×10^6) were co-transfected with 500 ng of pAcH6N1-Cre and 100 ng of *Bsu*36 I digested BacPAK6 viral DNA using the lipofectin reagent.

As shown in Figure 3.2, BacPAK6 DNA digested with *Bsu*36I releases two small fragments, one of which contains part of the ORF 1629 gene which is essential for viral replication. This provides two selection steps for the isolation of recombinant viruses: First, if the large fragment recircularises by itself, part of ORF 1629 is still missing and so the viral DNA cannot give rise to viable viruses. If however it recombines (in a double recombination event) with a transfer vector containing the missing sequences, the resulting circular viral DNA can produce viable viruses. Non-recombinant viral DNA cannot form lytic plaques because they do not have the essential gene for viral replication. Whereas single crossover recombinants in which the entire plasmid DNA has integrated into the viral genome are inherently unstable because they contain duplications of portions of the viral genome (i.e., the flanking regions present in the transfer plasmid DNA duplicate existing regions in the virus genome).

Figure 3.3b- Schematic representation of PCR designed to amplify cre gene from plasmid pMC1-Cre.A combination of oligonucleotide sequences designed as forward (CB1741) and reverse (CB1742) primers were used in a PCR reaction to amplify the 1.1 kb *cre* gene (shown as blank box) from pMC1-Cre. The translational initiation codon of cre is shown as ATG and the engineered nuclear localisation signal (NLS) is shown as a black box.

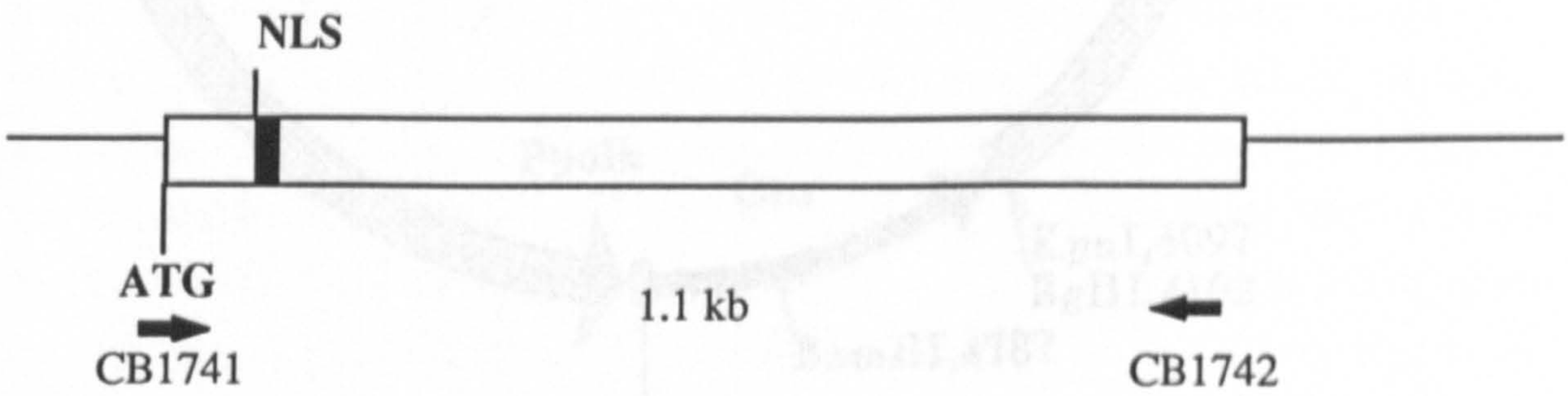
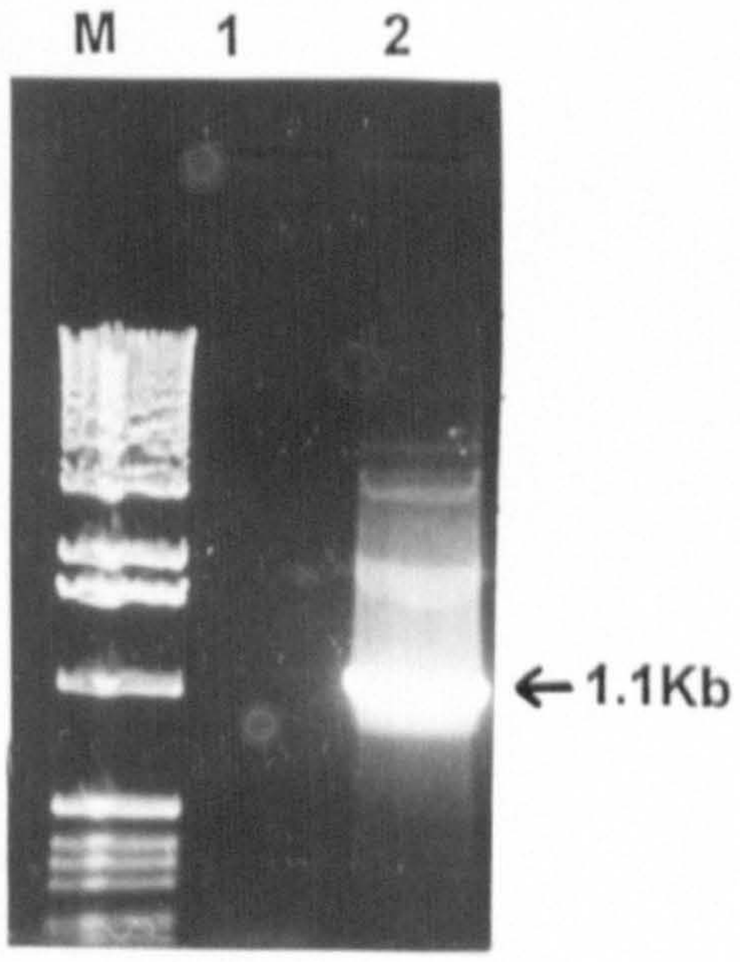


Figure 3.3c- Agarose gel electrophoresis analysis of PCR product.A PCR product of 1.1 kb was amplified from PCR using pMC1-Cre (lane 2). Lane 1 corresponds to primers only reaction (negative control). M indicates λ Hind III standard marker.



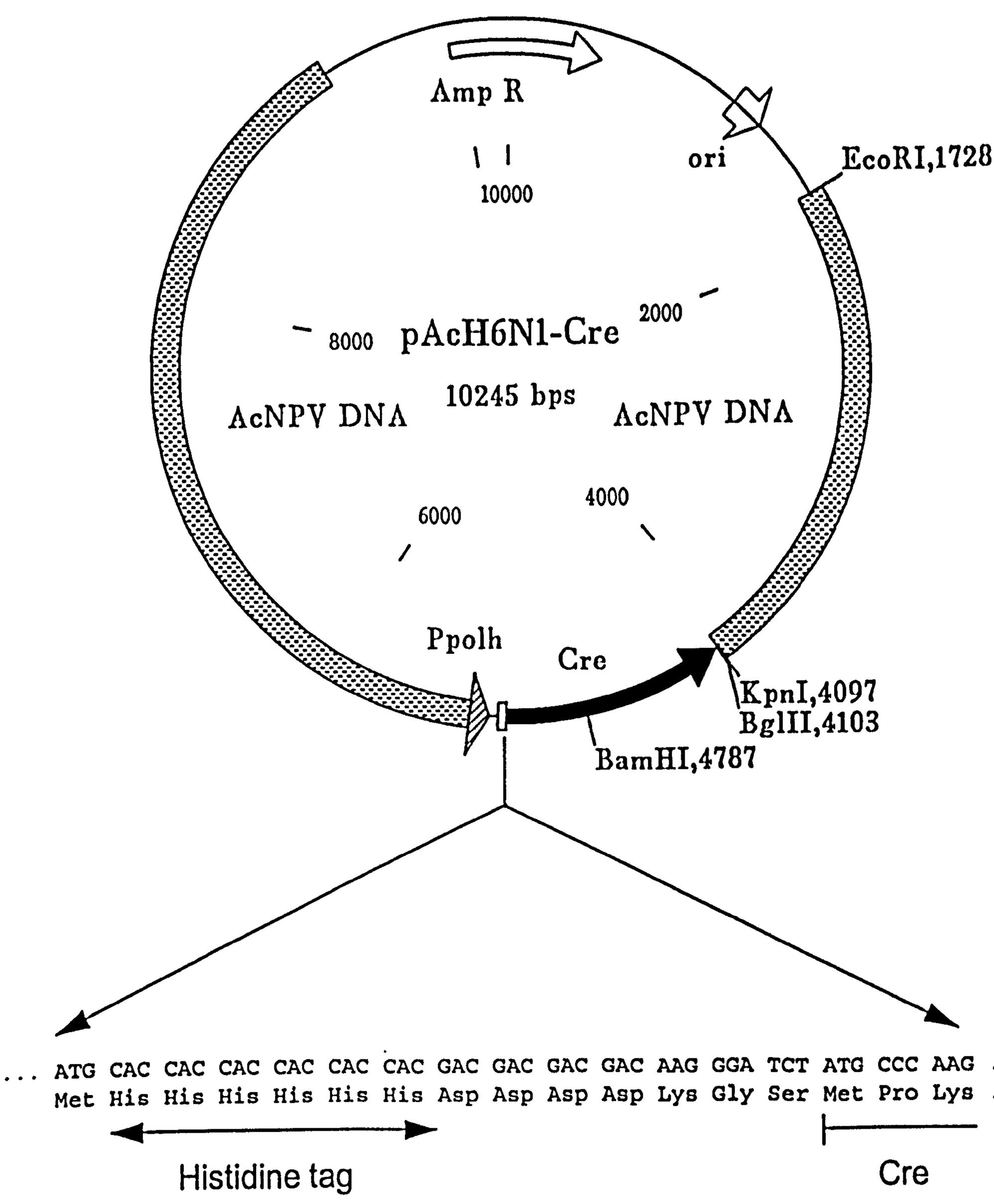


Figure 3.3d- Map of the chimeric transfer vector pAcH6N1-Cre

The chimeric transfer vector was constructed by cloning the recombinase gene (*cre*) from pMC1-Cre as a *Bgl*II fragment into the *Bam*HI site in pAcH6N1. The transfer vector, pAcH6N1 was derived from pVL1392 (Invitrogen). It contains 6 Kbp of the wild-type AcNPV genomic DNA (shown as dotted boxes), including the polyhedrin promoter, protein coding region and polyadenylation signal (see Figure 3.3a). In addition, it contains six histidines residues at position +35 with respect to mRNA transcriptional initiation site at position +1. The resulting vector, pAcH6N1-Cre contains the *cre* coding region fused in-frame to the upstream six histidine residues under the transcriptional control of the polyhedrin promoter, Ppolh (shown as slashed arrowhead).

Second, BacPAK6 contains a β-galactosidase (*lac*Z) marker gene which hydrolysis the chromogenic indicator X-gal to give blue plaques in a plaque assay. However, a double recombination event loses the *lac*Z gene. Therefore, double crossover recombinants can be distinguished from non-recombinant viral DNA or single crossover recombinants because they are phenotypically lacZ$^-$ (detected as colourless plaques in the presence of X-gal in a plaque assay).

Following transfection, 2 ml of supernatant from the infected cells putatively containing recombinant viruses were serially diluted in growth (TC100) medium to a million fold, and used to infect fresh plated Sf9 insect cells for plaque assay and purification. In subsequent plaque assays, colourless plaques were observed indicating that the double recombination reaction between the transfer vector pAcH6N1-Cre and the *Bsu*36I digested BacPAK6 viral DNA had occurred. Isolated plaques were picked for purification by resuspension in TC100 medium and serial dilutions used to re-infect insect cells for clonality. A final titre value of 4×10^7 plaques forming units (pfu)/ml was calculated for the purified plaques.

In order to verify further that the co-transfection of pAcH6N1-Cre and BacPAK6 viral DNA into insect cells generated recombinant viruses containing the *cre* gene, viral genomic DNA prepared from purified plaques were used in a PCR reaction. A combination of oligonucleotide sequences designed as forward primer (BK59-CGTGTTTTAATACGCCGGACC) to anneal within the polyhedrin locus upstream of *Kpn*I site at position 4967bp-4946bp in pAcH6N1 (see Figure 3.3d), and reverse primer (CB1741- GACAAGAGATCTATGCCCAAGAAGAAGAGGAAGG) to anneal at the 5' end of the *cre* gene, were used to amplify sequences viral genomic DNA. For each recombinant baculovirus analysed (lanes 3-8), a fragment of the expected 1.1 Kbp size was amplified (Figure 3.4). This fragment was the same size as the PCR product from the positive control, pAcH6N1-Cre (lane 9). In contrast, this 1.1 Kbp product was not observed in the negative control (wild-type baculovirus). Instead, products of higher molecular weights were observed suggesting non-specific DNA amplification from the wild-type viral genome.

This result provides further confirmation that the *cre* gene is present in the recombinant virus genome and that the cre gene is correctly orientated with respect to the viral promoter to allow for efficient expression.

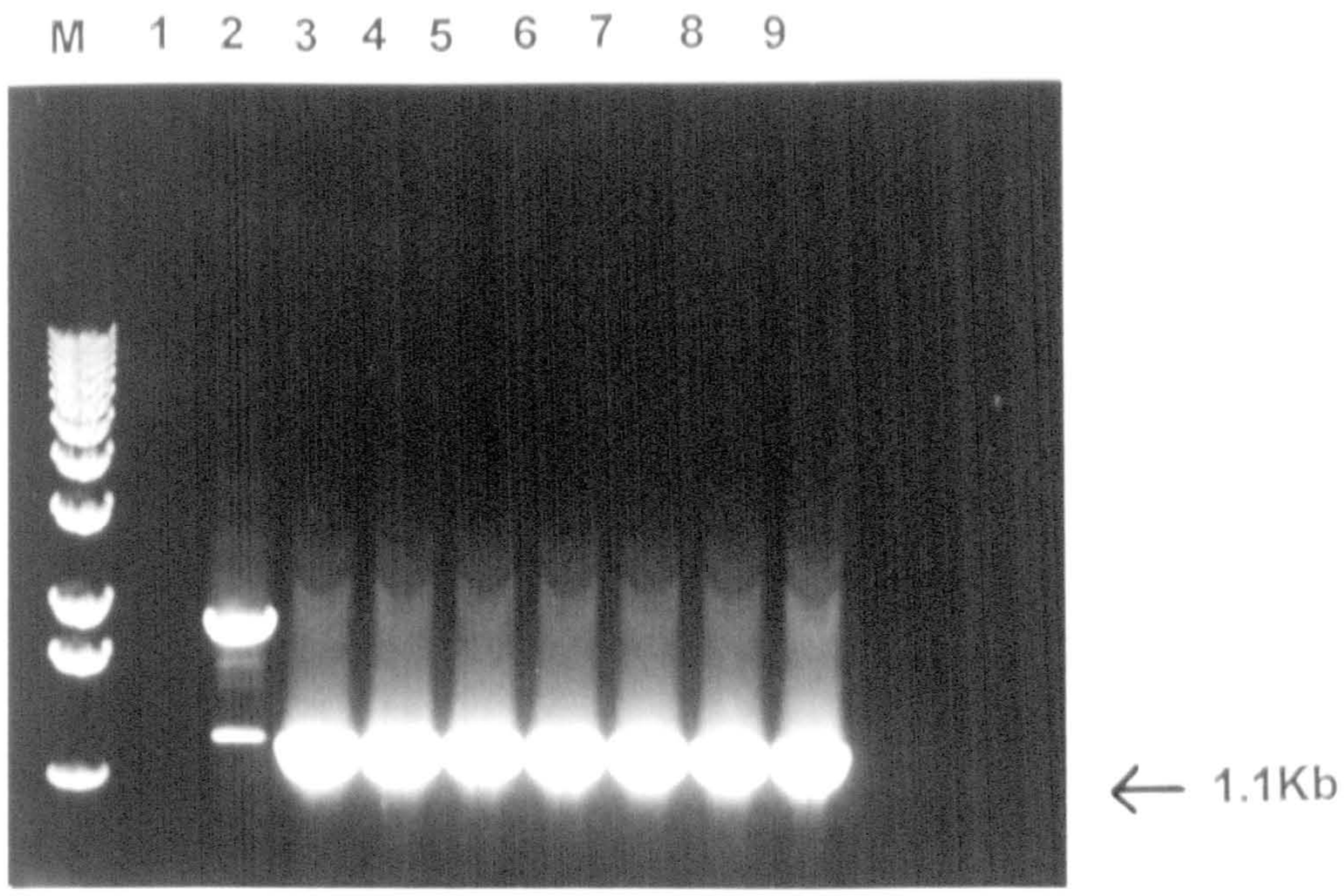


Figure 3.4- Agarose gel separation of PCR reactions of recombinant baculoviral genomic DNA.

PCR reactions were done using viral genomic DNA isolated from wild-type baculovirus (AcMNPV; as a negative control), the recombinant baculoviruses and the chimeric transfer vector, pAcH6N1-Cre (as a positive control). A pair of oligonucleotide primers designed to prime within the polyhedrin promoter and the *cre* gene was used to amplify sequence from the viral genomic DNA. In each reaction, a 1.1 Kbp fragment was amplified for the recombinant baculoviruses (lanes 3-8), the same size fragment as the positive control, pAcH6N1-Cre (lane 9). Different size fragments were observed in the negative control, AcMNPV (lane 2). Lane 1 corresponds to primers only reaction (negative control). M indicates λ Hind III standard marker.

3.5.3 Metabolic labelling of infected cells

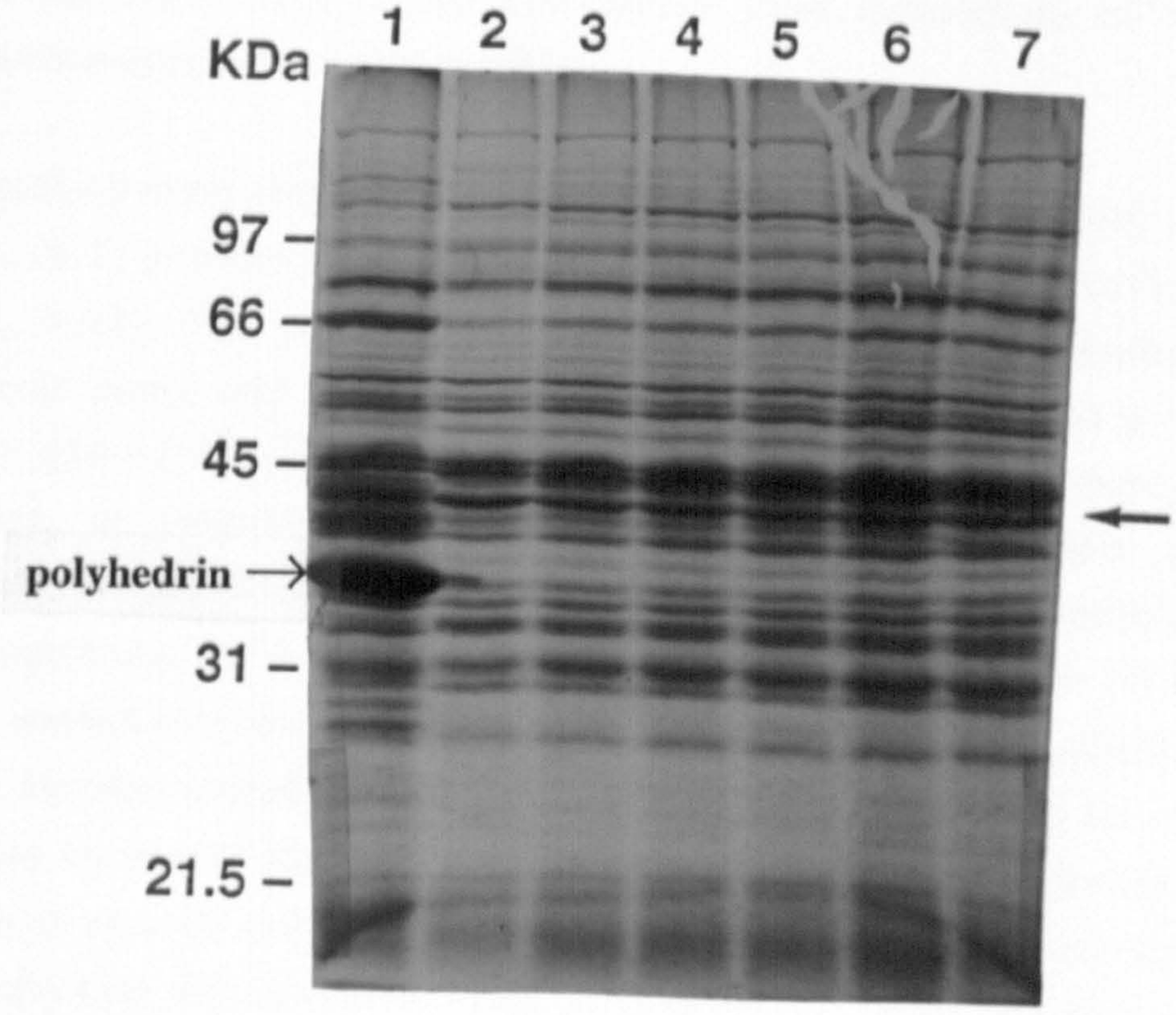
In order to evaluate whether the recombinant virus was producing Cre protein, and if so, to determine the levels of expression, Sf9 cells infected with the recombinant baculovirus were metabolically labelled with ^{35}S labelling mix and total infected cell proteins analysed by SDS-PAGE gels and Coomassie blue staining.

There are several approaches that can be used to evaluate whether a recombinant virus is producing the required protein. The simplest is the characterisation of whole-cell lysates by Coomasie blue staining or immunoblotting. Coomasie blue staining or immunoblotting of whole-cell lysates analysed on SDS-PAGE gels allows the detection of proteins during the infection process. If the cells are incubated with a radiolabelled amino acid for a short period prior to lysis, proteins expressed during the labelling period (i.e. very late infection) can be detected as long as they are relatively highly expressed. This is because the strength of a radioactive signal on an autoradiograph is proportional to the amount of radioactivity incorporated by a protein during the labelling period, which in turn, is partly dependent on the rate of synthesis of the protein.

Plaque-purified recombinant viruses or the wild-type baculovirus were used to infect Sf9 (2×10^6) cells at 10 pfu/cell, and total protein products analysed 72 hours post infection (h.p.i) by Trans ^{35}S-labelling (see section 2.6.2). Total cell lysates prepared after labelling, were analysed by SDS-PAGE and autoradiography (Figure 3.5).

In Sf9 insect cells infected with the wild-type virus an intense signal of ~33 KDa corresponding to the polyhedrin protein was readily detected on SDS-PAGE (lane 1). In cells infected with the recombinant virus, the signal corresponding to the polyhedrin protein was not observed. Instead a signal migrating at about 38 KDa (lanes 2 to 7) was observed which was absent in the wild-type virus control. This signal probably corresponds to the Cre protein. It was also noted that the intensity of these signals were much weaker than the the wild-type signal. If the 38 KDa signal corresponds to the Cre protein, then this shows that the expression of the Cre protein was much lower than the polyhedrin protein in the wild-type infection. This observation has been reported for the expression of other heterologous proteins (described in section 3.4.3).

Figure 3.5- The rate of Cre synthesis analysed on an SDS-PAGE.
S. frugiperda cells (2 × 106) were infected at 10 pfu/cell with the recombinant baculovirus or wild-type baculovirus, AcMNPV and labelled for 2 hours with Trans 35S-label mix at 72 h.p.i. Cells were lysed and aliquots analysed directly by SDS-PAGE and autoradiography. Lane 1; lysate from vAcMNPV infected cells, lanes 2 to 7; lysates from the recombinant baculovirus infected cells. The arrow (on the right) indicates a band of approximately 38 kDa corresponding to the Cre protein. A band corresponding to the wild type (polyhedrin) protein is shown on the left. Molecular weight standard (kDa) are indicated to the left of the gel.


KDa
1
2
3
4
5
6
7
97
66
45
polyhedrin
31
21.5

3.5.4 Purification of Cre protein from immobilised Ni^{++}-NTA spin column

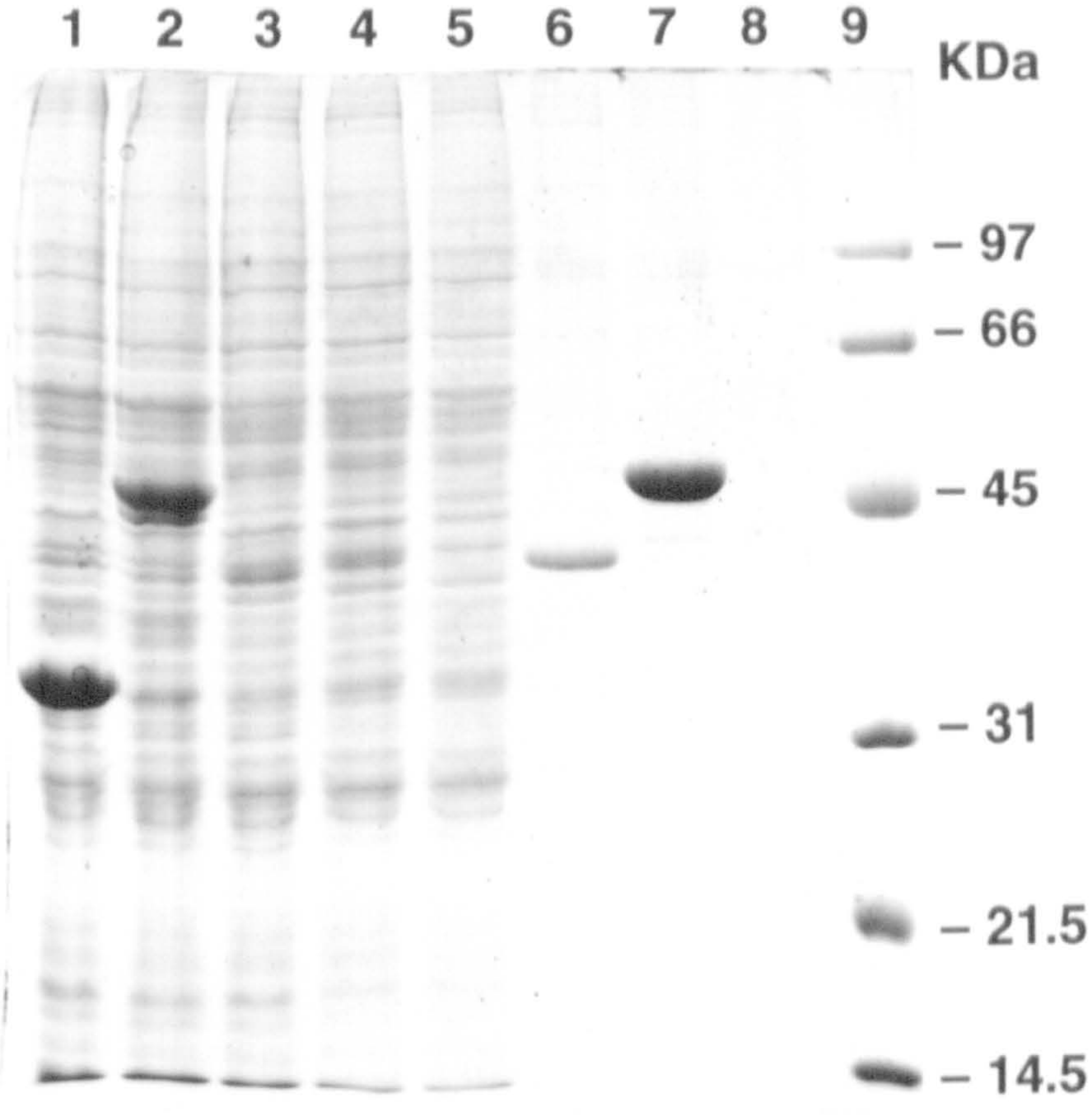
Although the baculovirus/insect system allows the high level expression of recombinant proteins, purification from host endogenous proteins can be a time-consuming task and a rate-limiting step in obtaining large quantities of material. To overcome this problem the Cre protein was expressed as a fusion with a hexa-histidine tag. Because histidine residues interact with a strong binding affinity for metal chelate adsorbents, a rapid single-step purification of the Cre protein from immobilised Ni^{++}-NTA affinity chromatography column was possible.

Immobilised metal chelate affinity chromatography, first used to purify protein (Porath *et al.*, 1975) provides a simple one-step method for rapid protein purification (Hochuli *et al.*, 1987). This purification system uses a metal chelate adsorbent, Ni^{++}-NTA (nitrilo-tri-acetic acid, charged with nickel ions) resin which has a strong binding affinity ($Kd=10^{-13}$, pH 8.0; far greater than the affinity between most antibodies and antigens, or enzymes and substrates) for proteins or peptides containing six consecutive histidine residues at either their N- or C-terminus (Hochuli *et al.*, 1988). This means that any host proteins that bind non-specifically to the NTA resin can be easily washed away under relatively stringent conditions, without affecting the binding of the histidine tagged proteins. Elution of tagged proteins from the column can be achieved by several methods. Reducing the pH will cause the histidine residues to become protonated and so dissociate from the Ni^{++}-NTA. Elution can also be achieved by competition with imidazole, which binds to the Ni^{++}-NTA and displaces the tagged proteins (Janknecht *et al.*, 1991). The presence of specific proteolytic cleavage site allows for the histidine tag moiety to be cleaved from the desired protein. However, the small size means that it does not affect the physical or functional properties of proteins, hence there is rarely any need to remove it from the recombinant protein after purification. Additionally, monoclonal antibodies can be raised against the tag moiety which allows the detection of proteins without the need to raise specific sera against the protein of interest (Zentgraf *et al.*, 1995).

Sf9 (5×10^9) cells were infected with the recombinant baculovirus at 10 pfu/cell, and cell extract prepared 72 h.p.i applied onto an immobilised Ni^{++}-NTA affinity chromatography column. The column was washed with 5 mM imidazole and bound proteins eluted with 50 mM imidazole. Samples taken after each purification step were analysed by SDS-PAGE and visualised by Coomasie blue staining (Figure 3.6). Cell extract prepared from pAcH6N1-E1.8 (the interferon protein E1.8 expressed as a

Figure 3.6- SDS-PAGE analysis of Cre purification from immobilised Ni^{++}-NTA spin column.

S. frugiperda cells (5 × 109) were infected at 10 pfu/cell with the wild-type baculovirus (AcMNPV; negative control), recombinant baculovirus expressing an interferon protein E1.8 (BacPAK6/E1.8; positive control), and a recombinant baculovirus expressing the Cre recombinase protein (BacPAK6/Cre). Cells were harvested 72 h.p.i by centrifugation and resuspended in binding buffer (described in materials and methods). The cell suspension was freeze-thawed to lyse the cells. Cellular debris was removed by centrifugation and proteins purified from the supernatant by affinity chromatography on Ni^{++}-NTA spin column. Samples taken after each purification step were analysed by SDS-PAGE and proteins visualised by Coomassie blue staining. Lane 1; total cell extract from AcMNPV infection, lane 2; total cell extract from BacPAK6/E1.8 infection, lane 3; total cell extract from BacPAK6/Cre infection, lane 4; supernatant from BacPAK6/Cre infection, lane 5; eluate from Ni^{++}-NTA column in 5 mM Imidazole (BacPAK6/Cre infection), lane 6; eluate from Ni^{++}-NTA column in 50 mM Imidazole (BacPAK6/Cre infection), lane 7; eluate from Ni^{++}-NTA column in 50 mM Imidazole (BacPAK6/E1.8 infection;), lane 8; eluate from Ni^{++}-NTA column in 50 mM Imidazole (AcMNPV infection; negative control), lane 9; standard protein markers. Molecular weight standards (kDa) are indicated to the right of the gel.


1
2
3
4
5
6
7
8
9
KDa
– 97
– 66
– 45
– 31
– 21.5
– 14.5

fusion gene with histidine tag) was used for positive control, whilst cell extract from the wild-type baculovirus infected cells was used for negative control.

The result shows that in the presence of 5 mM imidazole, no Cre protein was eluted from the Ni^{++}-NTA affinity column (lane 5), whereas in the presence of 50 mM imidazole, most of the Cre protein eluted from the column migrating as a single band of about 38 KDa on an SDS-PAGE (lane 6). As expected from the positive control, the E1.8 protein was eluted from the column in 50 mM imidazole (lane 7), whereas the wild-type polyhedrin protein which was not expressed as a fusion protein was eliminated in the wash through (lane 8). These results clearly show that the Cre protein expressed as a fusion protein with a hexa-histidine tag at its N-terminus can be purified in a one-step method from immobilised Ni^{++}-NTA affinity chromatography column. The absence of Cre protein in the wash through shows that there is a strong binding interaction between the histidine residues and the Ni^{++} ions.

After the purification of Cre protein from the baculovirus/insect expression system, it was important to establish that the protein is biologically active. In order to do this, a simple *in vitro* assay system was set up using as a substrate, plasmid pAT-lacZloxP2. The construction of this plasmid is described in the next section.

3.6 Construction of pAT-lacZloxP2

In the review of the Cre-loxP recombination reaction described in the introductory chapter, it was mentioned that native Cre recombinase catalyses site-specific recombination in plasmid substrates containing loxP sites *in vitro* in the presence of Mg^{2+} without any external energy cofactors (Abremski *et al.*, 1983). Although these plasmids were available, a substrate plasmid which contained the loxP sequences in enzyme recognition sites that would be easy to manipulate by substitution with other sequences (described fully in chapter 4) and which can easily be assayed was required. For these reasons, pAT-lacZloxP2 was constructed.

Plasmid pAT-lacZloxP2 is a 4.0 Kbp pAT153-derived vector containing β-lactamase (ampicillin resistance [amp^r]) gene and pBR322 origin of replication (ori). The plasmid was constructed in the following steps:

(i) The multiple cloning site (m.c.s) within the α peptide sequence of the *lac*Z (lacZ-α) gene was removed by digesting pBS KS (-) with *Bss*HII. The *Bss*HII cohesive ends

were trimmed by digestion with SI nuclease followed by blunt end ligation to generate ΔpBS KS(-).

(ii) A combination of oligonucleotide sequences designed as forward primer (B142- gacgtcGAATTCTTTCCATTCGCCATTCAGGCT) to anneal the 5' end of the *lac*Z-α gene and the reverse primer (B143- gacgtcaAGCTTGCGCAACGCAATTAATGTGA) to anneal the 3' end of the *lac*Z-α gene were used to amplify the *lac*Z-α gene from Δ pBS KS(-) in a PCR reaction (the introduced *Eco*RI site in B142 and *Hind*III site in B143 are underlined and lower case letters indicate non-specific nucleotide extensions). The PCR product was analysed by agarose gel electrophoresis and subsequently purified and digested with *Hind*III and *Eco*RI restriction endonucleases.

(iii) The lacZ-α gene was cloned into the unique *Hind*III and *Eco*RI restriction endonuclease sites in pAT153 vector to generate pAT-lacZ.

(iv) Oligonucleotides corresponding to the loxP sites were synthesised as complementary strands with the necessary protruding ends required for cloning. The plasmid pAT-lacZloxP was generated by inserting a loxP site into the *Eco*RI site in pAT-lacZ. A second loxP site was inserted into the *Hind*III site in pAT-lacZloxP to generate pAT-lacZloxP2 (Figure 3.7). The orientation and sequence of the loxP sites and the *lac*Z-α gene in the plasmid were validated by restriction endonuclease digestions and by sequencing.

The resulting plasmid, designated pAT-lacZloxP2, contains two loxP sites flanking the α peptide of *lac*Z gene. The loxP sites are directly repeated relative to each other, so that Cre-mediated recombination results in the excision of the *lac*Z gene flanked by the two catalytic sites. This is described in the next section.

3.7 Analysing the biological activity of the recombinant Cre protein

In order to investigate if the baculovirus/insect derived Cre protein could mediate site-specific recombination between loxP sites without any external energy co-factor(s) such as ATP (a characteristic of native Cre recombinase), a simple *in vitro* assay system (see section 2.8.1) containing pAT-lacZloxP2 in a solution of Mg^{2+}, Na^{+} and BSA was used. As described in chapter 4, the presence of these cofactors in the reaction mixture optimises the efficiency of the Cre reaction, however they are not absolutely essential because the omission of one or two of these cofactors does not prevent recombination.

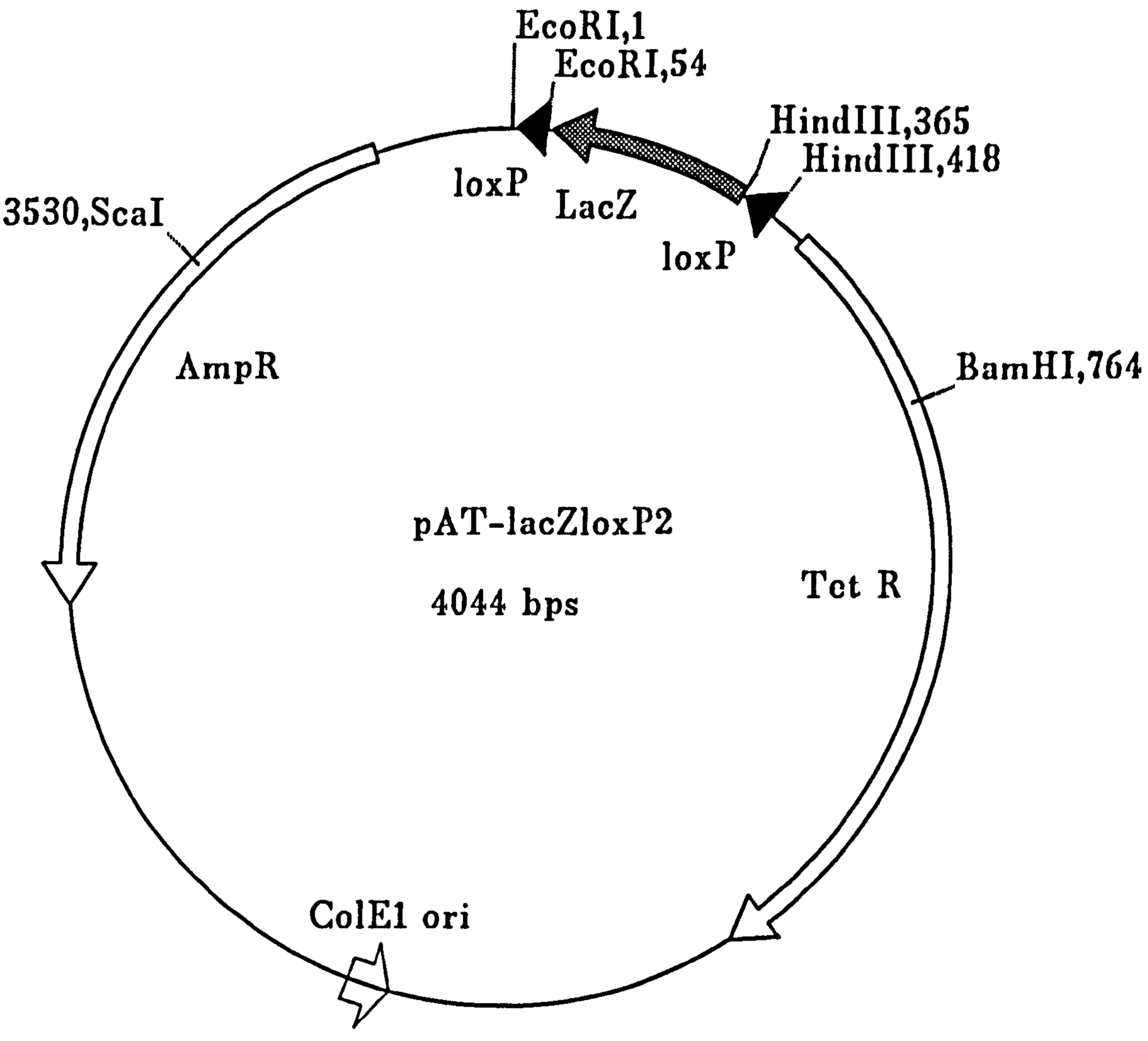


Figure 3.7- Map of plasmid pAT-lacZloxP2

pAT-lacZloxP2 is a 4.0 Kbp pAT153-derived vector containing the β-lactamase ampicillin resistance (amp^r) gene and pBR322 ori (origin of replication). The α peptide of *lac*Z was cloned into the unique *Hind*III and *Eco*RI sites in pAT153. Oligonucleotide sequences comprising loxP sites were cloned sequentially into the *Hind* III and *Eco*R I sites in the recombinant vector. The resulting plasmid contains two loxP sites in direct repeats flanking the α peptide of *lac*Z gene.

Cre-mediated recombination between the loxP sites leads to the excision of the *lac*Z gene, giving rise to two products: the vector backbone containing a loxP site (product A) and the *lac*Z gene also containing a loxP site (product B). (These fragments then become substrates themselves because each of them contains an active loxP site). Thus the Cre protein also catalyses site-specific recombination between the two products converting them back to the original substrate or to other recombination intermediates such as multiple integrations which are not detected here because they occur at very low frequency or are non-viable in cells. However, the integration event in general is less efficient than the excision event (Figure 3.8). The molecular species generated by the recombination event can be analysed either directly by restriction endonuclease digestion or by the phenotypic colour selection of substrates and products following an *in vivo* assay described in the following sections.

3.7.1 Restriction endonuclease digestion analysis of Cre-mediated recombination

Plasmid pAT-lacZloxP2 was incubated with various amounts of the purified recombinant Cre protein and the DNA was subsequently purifed and digested with *Bam*HI and *Sca*I and analysed by agarose gel electrophoresis (Figure 3.9). Restriction endonuclease digestion of pAT-lacZloxP2 with *Bam*HI and *Sca*I yields two fragments of 2.7 and 1.3 Kbp. After Cre mediated excision of the insert between the loxP sites, restriction digestion results in the formation of a 0.9 Kbp fragment in addition to the 2.7 and 1.3 Kbp.

The occurrence of the 0.9 Kbp band in addition to the 2.7 and the 1.3 Kbp fragments indicates that the Cre protein has recombinase activity. However, because the Cre recombinase activity is reversible it is unlikely to observe a complete excision of the insert fragment. As described in section 4.5, the efficiency of the excision event can be quantified by counting hybridisation signals from a Southern blot analysis using a phosphoimager. An alternative method of determining the recombination efficiency is described in section 3.7.2.

3.7.2 Analysing the results of Cre-mediated recombination *in-vivo*

If recombination is genuinely site-specific at the loxP sites and mediated by the Cre protein, then the consequence of the recombination event should be the excision of the *lac*Z-α gene resulting in recombinant products which are phenotypically LACZ$^-$. The

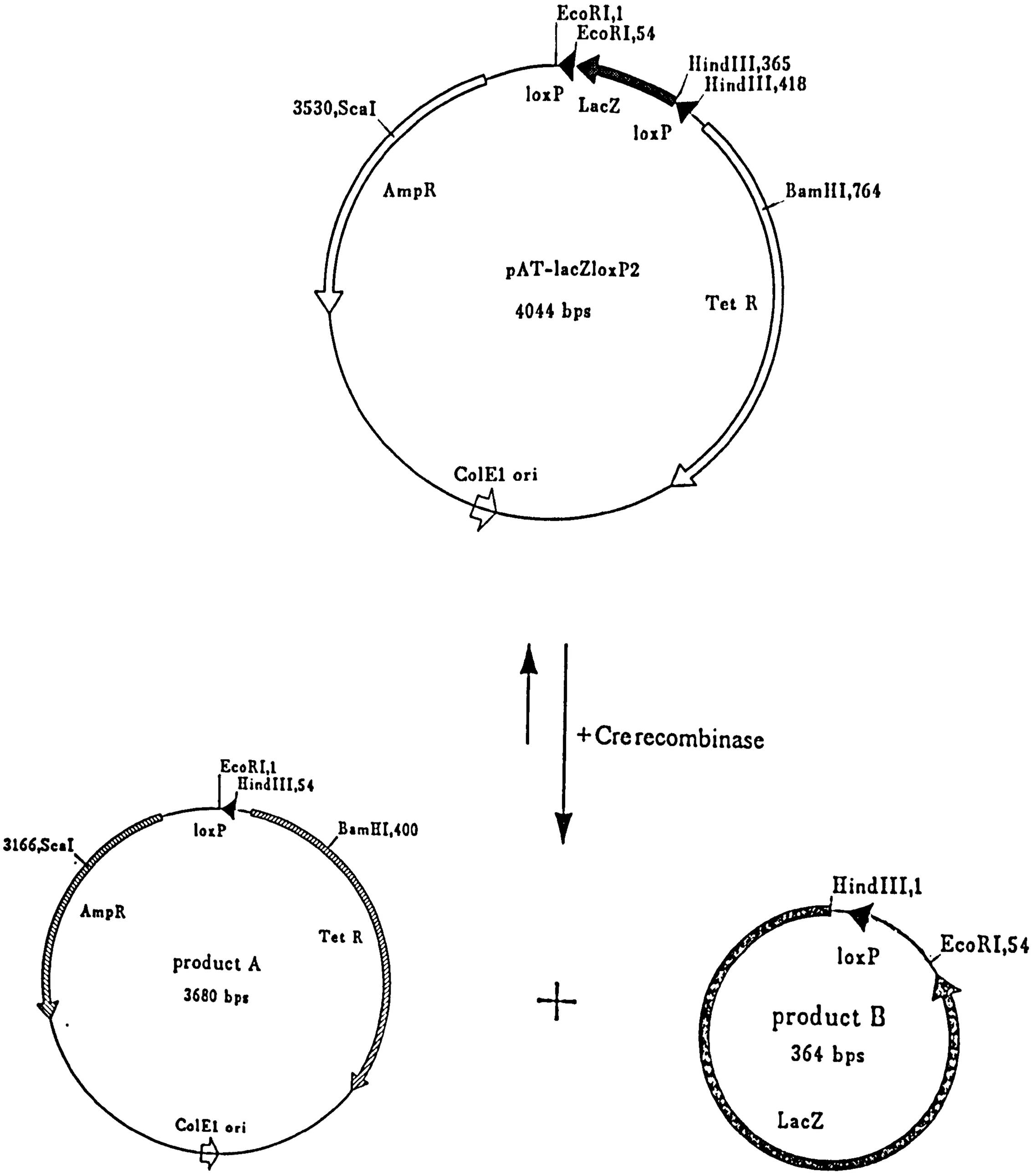


Figure 3.8- Diagrammatic presentation of Cre-mediated excision reaction.
Cre mediated site-specific recombination at loxP sites in the plasmid, pAT-lacZloxP2 excises the intervening *lac*Z gene. The event gives rise to two products, the vector backbone (product A) and the *lac*Z gene (product B). Each product contains a loxP site and is itself a substrate for Cre mediated integration reaction.

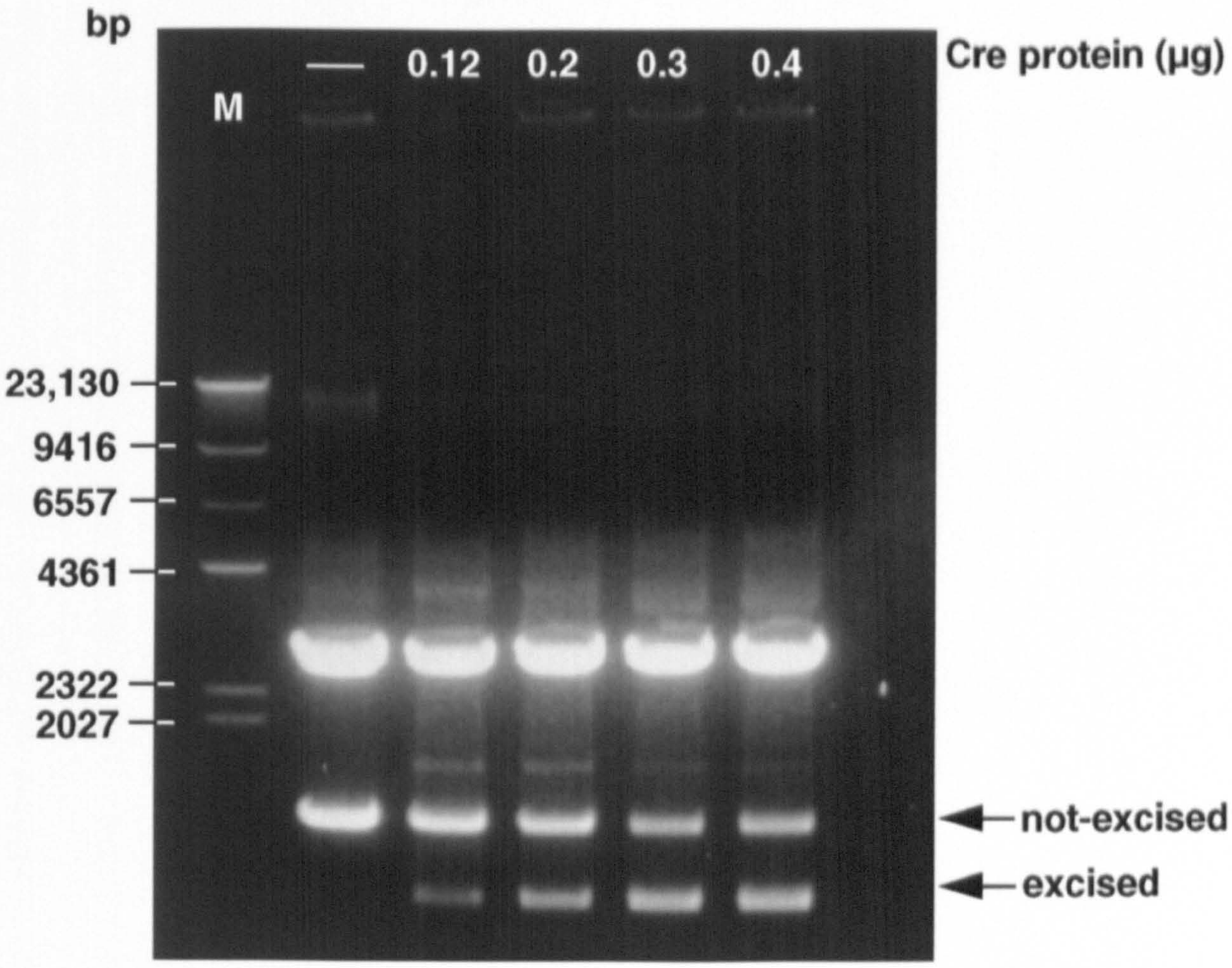


Figure 3.9- Restriction digestion analysis of the Cre-mediated recombination.

The plasmid pAT-lacZloxP2 (2 µg) was incubated with various amounts of the purified Cre protein in a 30 µl reaction volume containing 50 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$; 30 mM NaCl; 1 mg bovine serum albumin/ml for 20 minutes at 37°C. The reactions were purified by phenol/chloroform extraction followed by ethanol precipitation. Plasmid DNA was subsequently digested with *Bam*HI and *Sca*I and separated on 0.7% agarose gel. DNA was visualised by ethidium bromide staining.

recombination product containing the *lac*Z-α gene does not contain an origin of replication (ori) and will therefore not replicate in cells, whereas the product that is resistant to ampicillin contains an ori and will therefore replicate in cells.

Aliquots of the recombination reaction (containing ~50 ng of DNA) were transfected into *E. coli* DH5α competent cells and dilutions plated on L-agar (containing ampicillin, X-gal and IPTG). In the absence of Cre protein only blue colonies were observed on L-agar plates. However in the presence of the recombinant Cre protein both blue and white colonies were observed indicating that Cre-mediated recombination gave rise to products that are phenotypically LACZ$^-$. The recombination event can therefore be assayed semi-quantitatively *in vivo* by the phenotypic colour selection of substrates and products on L-agar plates containing X-gal because the substrate pAT-lacZloxP2 contains the *lacZ* gene whose protein product, β-galactosidase hydrolyses the chromogenic indicator X-gal to a blue phenotype, whilst the product containing the ampicillin resistant gene cannot hydrolyse X-gal.

3.7.3 Sequence analysis of excision products

If the Cre-mediated excision reaction gives rise to products which then becomes substrates for the reverse integration process, then an active loxP site must be present in each recombination product. In order to investigate this, DNA prepared from five white colonies (corresponding to a recombination product described above), were used in sequencing reactions. The reactions were analysed by polyacrylamide gel electrophoresis and visualised by autoradiography (Figure 3.10). Results showed that recombination event gave rise to products that contained an intact loxP site in all the white colonies analysed. The presence of active loxP sites in the recombination products demonstrates that the reversible reaction can occur.

The quantitative analysis of the Cre-mediated excision reaction described in this section suggests that this assay system can be used to investigate various parameters of the Cre-loxP excision reaction (described in chapter 4). In order to do this, it was important to establish how much of the protein preparation is active Cre, for this reason the specific activity of the Cre protein was determined.

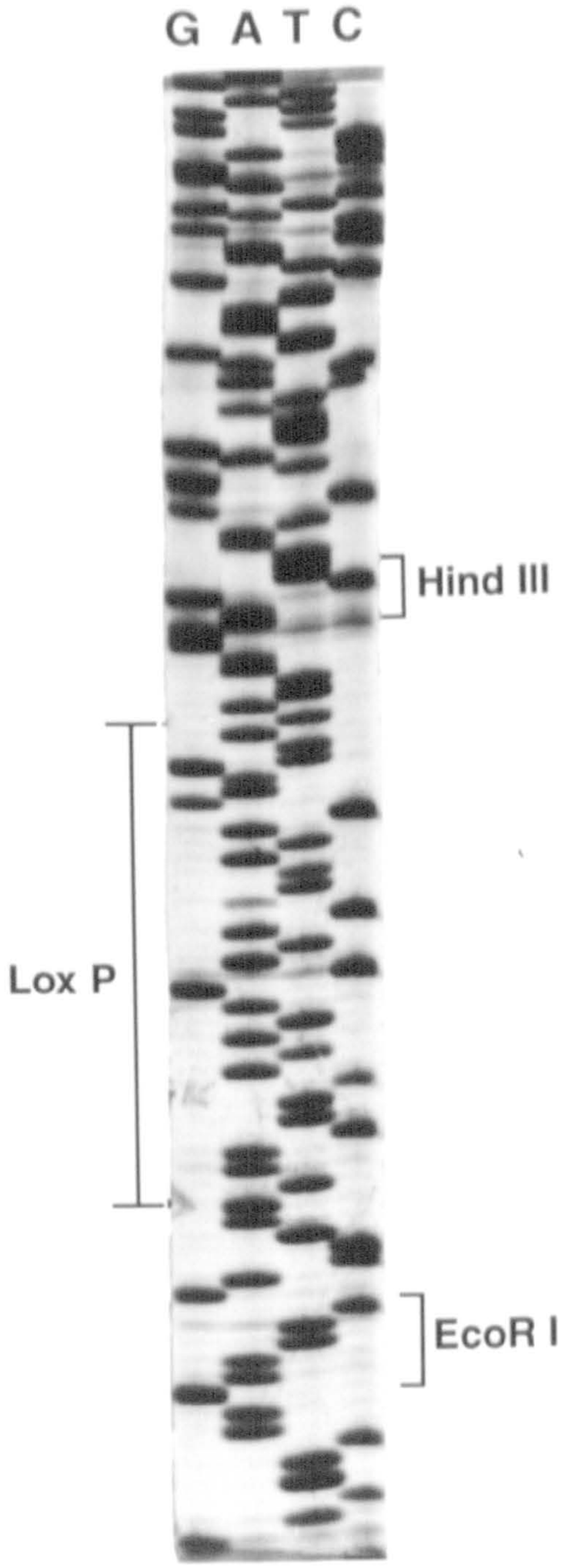


Figure 3.10- Investigating the recombination product by sequence analysis

Plasmid DNA prepared from white colonies (product A; see Figure 3.8) as a result of Cre-mediated recombination were sequenced using an oligonucleotide sequence which primes within the amp^r gene in pAT-lacZloxP2 (see Figure 3.8). The sequencing reaction was analysed by polyacrylamide gel electrophoresis and visualised by autoradiography. The sequence analysis showed a loxP site (consisting of two 13 bp inverted repeats [shown in boxes with arrow heads] flanking an 8 bp spacer region) and *Eco*RI and *Hind*III restriction endonuclease sites.

3.8 Determination of the specific activity of the Cre protein

The effect of varying the concentration of the recombinant Cre protein on the efficiency of site-specific recombination was investigated in order to determine the activity and the purity of the recombinant Cre protein.

The amount of Cre enzyme present or used in the recombination process is difficult to determine in absolute terms (e.g. grams) as its purity is often low and a proportion may be in an inactive, or partially active state. A more relevant parameter is the activity of the enzyme preparation, measured in terms of the activity unit (U), which is defined as the minimum amount of Cre needed to produce a maximum level of loxP X loxP intramolecular recombination under the assay conditions described in Abremski *et al.*, (1983). However the assay conditions which refer to optimal conditions, especially with regards to pH, ionic strength, temperature and substrate concentration may differ between laboratories. Therefore great care has to be taken over the consideration of these factors in defining the activity of Cre protein from different preparations.

The specific activity of the purified Cre protein was defined by incubating 2 µg of the substrate plasmid pAT-lacZloxP2 with different amounts of the Cre protein under standard *in vitro* assay conditions (described in section 2.7.4). After termination of the reaction, aliquots (containing ~50 ng of DNA) were transfected into *E. coli* DH5α and dilutions plated onto L-agar plates (as described in section 3.7.2). The results of the recombination, observed as white and blue colonies for recombinants and substrates respectively were quantified. The efficiency of excision (measured as % excision) is expressed as a measure of the ratio of white to blue colonies. In this experiment, a maximum excision was detected in the presence of 12 µg/ml of the Cre protein (afterwards the % excision "levelled off") per 2 µg of the plasmid (Figure 3.11). This corresponds to a specific activity of 27,000 U/mg (according to the calculation of Abremski and Hoess, 1984).

Thus, the specific activity of the Cre protein purified from the insect/baculovirus system is somewhat lower, although of the same order of magnitude with the specific activity for the native Cre protein (40,000 U/mg; Abremski and Hoess, 1984).

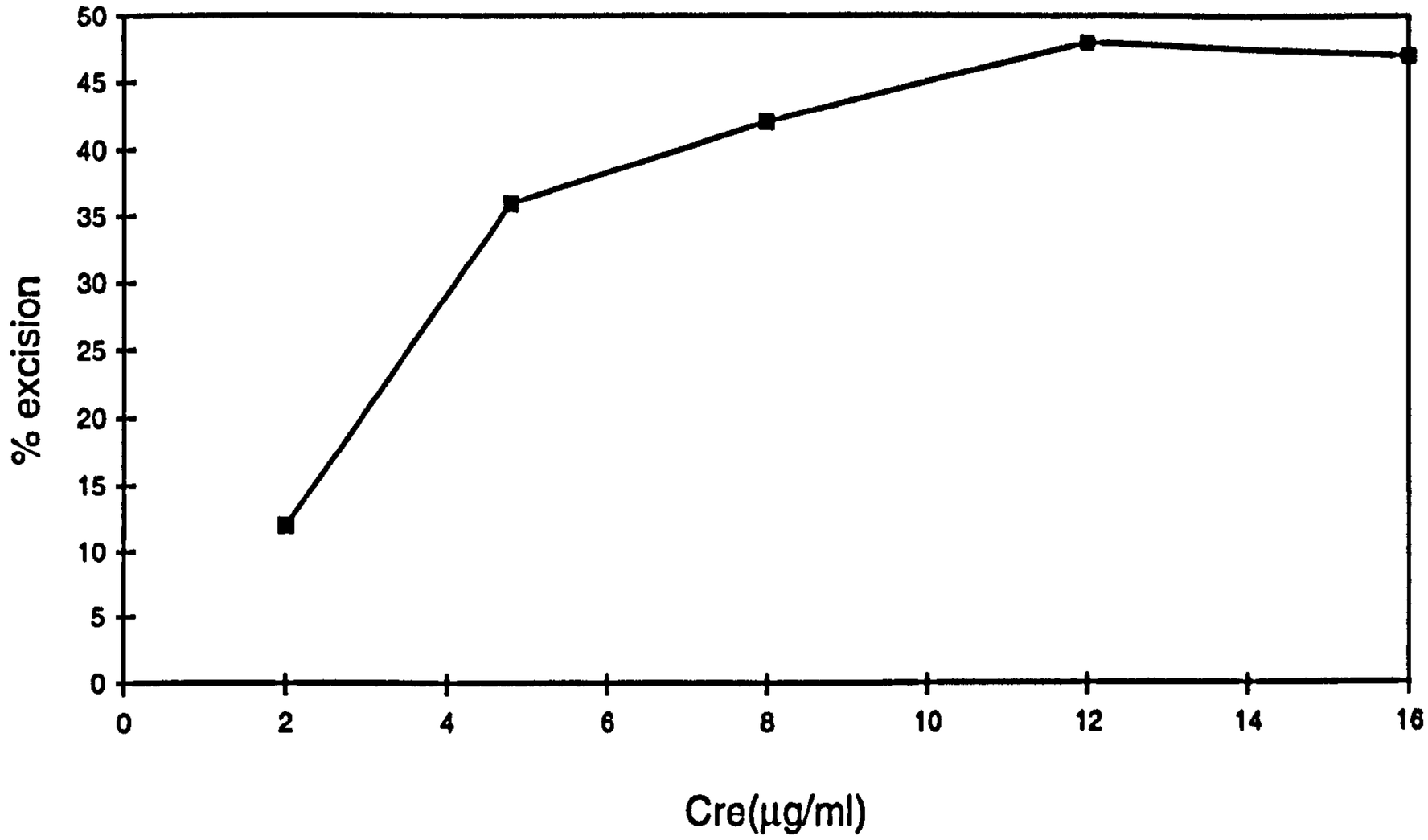


Figure 3.11- Determining the specific activity of the recombinant Cre protein.

Plasmid pAT-lacZloxP2 was incubated with 0.06, 0.12, 0.2, 0.3, or 0.4 µg of Cre protein in 30 µl reaction volume containing 50 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$; 30 mM Nacl; 1 mg bovine serum albumin/ml at 37°C for 20 minutes. After termination of each reaction, aliquots (containing ~ 50 ng of DNA) were transfected into *E. coli* DH5α competent cells and dilutions analysed on L-agar plates (containing ampicillin, X-gal and IPTG). The ratio of white white and blue colonies corresponding to recombinants and substrates respectively was expressed as a measure of the percentage (%) excision. The efficiency of recombination (% excision) is displayed as a function of the concentration (µg/ml) of the purified Cre protein in the reaction.

3.9 Summary

In this chapter, the Cre recombinase of bacteriophage P1 has been expressed in *Spodoptera frugiperda* insect cells using a recombinant baculovirus vector. The product was synthesised as a fusion protein with six histidine residues at the amino terminus allowing rapid and efficient purification in a single affinity chromatography step. Approximately 0.9 mg of active Cre protein was purified from 5×10^9 infected insect cells. This amount compares favourably with that achieved from *E. coli* source (0.4 mg/g of cell culture equivalent to ~ 2×10^{11} cells) Abremski and Hoess (1984). The baculoviral/insect derived Cre protein has the following features in common with native Cre protein: (i) a molecular weight of approximately 38 KDa, (ii) migrates as a single polypeptide band on an SDS-polyacrylamide gel, and (iii) catalysis of site-specific recombination in plasmid DNA containing loxP sites *in vitro*. The ease with which the protein can be expressed, purified and stored, suggests that the baculovirus/insect expression system can provide an alternative system for the large scale production of biologically active Cre protein that is required for further investigation into the mechanisms of the Cre-loxP reaction (described in chapter 4), and which can be directly introduced into mammalian cells to mediate targeted integration (described in chapter 5).

CHAPTER 4

Investigating the mechanism of Cre-mediated recombination *in vitro*

4.1 Introduction

In chapter 3, the expression of Cre protein in insect cells infected with recombinant baculovirus was described. Having established that the Cre protein purified from the baculovirus/insect system was biologically active, it is then possible to investigate the kinetic parameters of Cre-loxP site-specific recombination reaction. In this chapter, the Cre-loxP reaction was investigated *in vitro* with the aim of gaining a greater understanding into the mechanism of the reaction and also to identify irreversible substrates that could improve the efficiency of targeted integration in mammalian cells. As this chapter investigates the recombinant Cre protein's interaction with loxP sequences *in vitro*, a summary of the mechanism and factors that could influence Cre-loxP site-specific recombination is described below.

The first step on the pathway to site-specific recombination is the recognition of the loxP binding site by Cre protein. Although the mechanism by which Cre protein specifically recognises its loxP binding site is not known, the recognition of a specific DNA sequence by a protein has been studied for various proteins including the bacteriophage lambda Cro (Anderson *et al.*, 1981), *Salmonella* Hin recombinase (Feng *et al.*, 1994) and *E. coli* (F episome) γδ Resolvase (Yang and Steitz, 1995). Structural, biochemical and molecular genetic studies suggests that all characterised sequence-specific DNA binding proteins recognises DNA helix (Watson and Crick, 1953) and binds without disturbing "base-pairing". This is made possible because the Watson and Crick model proposed for DNA structure (known as B-DNA helix) suggests that each DNA helix consists of ten base-pair residues and each base pair is exposed in two separate "grooves" called major and minor grooves which forms the contact site for many proteins. As previously mentioned, a loxP site is made up of two 13 bp inverted repeat flanking an 8 bp spacer region and each 13 bp inverted repeat of a loxP site forms a binding domain for Cre protein. There are a number of ways for Cre protein to interact with its recognition loxP sites, however if the conformation of the loxP site is to fit the Watson-Crick model (1953), then there is within each 13-bp inverted repeat of the loxP site the necessary major and minor grooves required for interaction with the Cre protein.

The interaction of loxP sequences with Cre may involve hydrogen bonding between amino acids at the binding sites and also van der Waals interactions between side chains and base pairs. Protein-protein interactions may also help Cre hold onto the loxP more tightly when they bind through co-operative (synergistic) binding. These specific interactions, together with backbone salt bridges interactions may stabilise the Cre-loxP complex formation. Furthermore, a DNA molecule containing loxP sequence in solution is not a single static structure. Instead, there is a dynamic flux of charges that influences its topology in addition to the topological effects by other DNA interacting proteins. Cations such as Na^+, Mg^{2+} and polyamines (e.g. spermidine) neutralise DNA phosphate charges by forming cationic bridges with adjacent helices (Widom and Baldwin, 1980; Krasnow and Cozzarelli, 1982). As Cre protein binds loxP site(s), these neutralising charges must be displaced, as such, the equilibrium binding constant will depend not only on the concentration of the Cre protein and loxP, but also on the concentration of the charge-neutralising ion. Thus, the affinity of a Cre protein for loxP site can vary markedly with salt concentration.

There are several ways of investigating protein-DNA interactions. One approach describes the protection of DNA from nuclease digestion by protein. The position and sequence of the protected region can readily be identified by a technique called footprinting. In this approach, one end of DNA is first radiolabelled, protein is then added to the labelled DNA and the complex digested with DNase I. A control reaction is done where the DNA (without protein) is digested with DNase I. The resulting DNA fragments are separated according to size by gel electrophoresis. The gel pattern reveals a series of bands present in the control sample which is absent from the reaction containing the protein. These bands are missing because the protein shields the DNA from cleavage. The missing bands in the gel pattern identify the binding site on DNA (von Hippel *et al.*, 1984). Using this approach, Hoess *et al.*, (1984) showed that Cre recombinase binds a 34 bp sequence corresponding to the loxP site (see section 1.3.4). DNase I footprinting has the advantage of allowing the detection of protein binding DNA sequences, however it does not allow the detection of the affinity of protein for DNA binding sites or the determination of the stoichiometry of the protein-DNA interaction.

Another approach for studying protein-DNA interaction is known as band shift. A band shift (Revzin) assay results in the separation of free DNA and DNA-protein complexes according to their different electrophoretic mobilities which are resolved as discrete bands by polyacrylamide gel electrophoresis. Band shift assay has two

requirements; (i) that the complexes formed are stable and remain intact during the gel run, and (ii) that the complexes migrate at a rate different to that of the uncomplexed DNA fragment. This approach has the advantage of the determination of the stoichiometric ratio of protein specifically bound to DNA (Fried and Crothers, 1981; Garner and Revzin, 1981). Using this approach, Mack *et al.*, (1992) determined the absolute stoichiometry of the Cre-loxP complex as one molecule of Cre per inverted repeat, or two molecules per loxP site (see section 1.3.4).

As described in chapter 1, one round of site-specific recombination between two Cre-loxP complexes requires the breakage and reformation of four phosphodiester bonds (Jayaram, 1994). The reaction is completed by the combined action of four monomers of the recombinase binding at the recombination complex where strand cleavage and exchange occurs within the spacer region via a Holliday junction (Jayaram, 1994). The enzymology of the Cre-loxP reaction is very simple; Cre recombinase is the only protein required for strand breakage, exchange and resolution during site-specific recombination at loxP sites *in vitro*. No external energy cofactors (e.g. ATP) or accessory recombination proteins such as the *E. coli* recA are needed. Furthermore the *in vitro* recombination reaction can take place in low concentrations of a variety of buffers between pH 5 and 9 and ionic strength conditions between 50 and 350 mM (Abremski *et al.*, 1983; Sternberg and Hamilton, 1981).

The simplicity of the Cre-loxP site-specific recombination system has a major drawback for targeted integration events, in that recombination event is freely reversible. In site-specific DNA integration, recombination between a site on a circular molecule and a site on a plasmid or chromosomal DNA results in the insertion of the circular DNA into the plasmid or the chromosome. However the integrated DNA is flanked by two recombination sites of the same orientation and can therefore excise the inserted circular DNA if the recombinase is present. Because intramolecular excision is kinetically favoured over integration, insertion products are inherently unstable in the presence of the recombinase enzyme. In yeast and mammalian cells, a successful strategy used to obtain stable insertion products was to curtail post-integration recombinase activity through the transient provision of Cre recombinase (Baubonis and Sauer, 1993; Fukushige and Sauer, 1992; O'Gorman *et al.*, 1991; Sauer and Henderson, 1990). A second attempted strategy was to introduce selective base changes into recombination sites such that the product sites of a recombination reaction are less likely to undergo further recombination (Henrik *et al.*, 1995; Senecoff *et al.*, 1988). Further investigation into the mechanism of the Cre-loxP reaction may provide more information about aspects of the reaction that can be manipulated to

improve the efficiency of recombination which forms the basis of the work described in this chapter.

In this chapter the Cre-loxP reaction was investigated *in vitro* using two different approaches: First, Cre protein was used in gel retardation analysis with oligonucleotide sequences corresponding to loxP site, Flp recombinase recognition binding (frt) site, or a hybrid of loxP and frt half sites to investigate the specificity of the Cre protein for its recognition loxP site. As previously described in chapter 1, Cre and Flp belongs to the int family of recombinases and are particularly closely related to each other functionally. It was therefore of interest to investigate if Cre binds frt in order to establish novel Cre binding sites that could be used to improve Cre-mediated targeted integration. In addition, the stability of the Cre-loxP complex was investigated in competition assays and the effect of different incubation temperatures on the interaction.

Second, the Cre protein was used to investigate intramolecular (excision) reaction between two wild-type loxP sites or between a wild-type loxP and hybrid loxP/frt sites in the hope of identifying irreversible substrate that could be used to improve the efficiency of targeted integration. In addition, various factors such as temperature and different salt concentrations that could influence the efficiency of the Cre-mediated excision reaction were investigated. It was also important to establish that Cre protein could mediate intermolecular (integration) recombination between DNA fragments each containing a loxP site. This is because one of the main objectives of this thesis (described in chapter 5) was to introduce Cre protein directly into mammalian cells to mediate targeted integration. Finally the biochemical properties of the recombinant Cre protein reaction investigated here are compared with that of native Cre protein and aspects of the recombination process that can be manipulated to improve the efficiency of the recombination process are discussed.

4.2 Results

The following sections describes the results of the *in vitro* analysis of the Cre-loxP site-specific recombination reaction.

4.3 Recognition binding interaction using band shift assay

As previously described in the introduction to this chapter, band shift assays provides a the method for studying protein-DNA interaction *in vitro*. In this section the recombinant Cre protein was used in band shift assays with radiolabelled oligonucleotide sequences corresponding to either a loxP, frt or a hybrid of loxP/frt half sites. The binding reactions were done as described in section 2.7.2.

4.3.1 Investigating Cre-loxP interaction

In order to investigate if the recombinant Cre protein can recognise and bind to radiolabelled oligonucleotide sequence corresponding to a loxP site, a band shift assay was done.

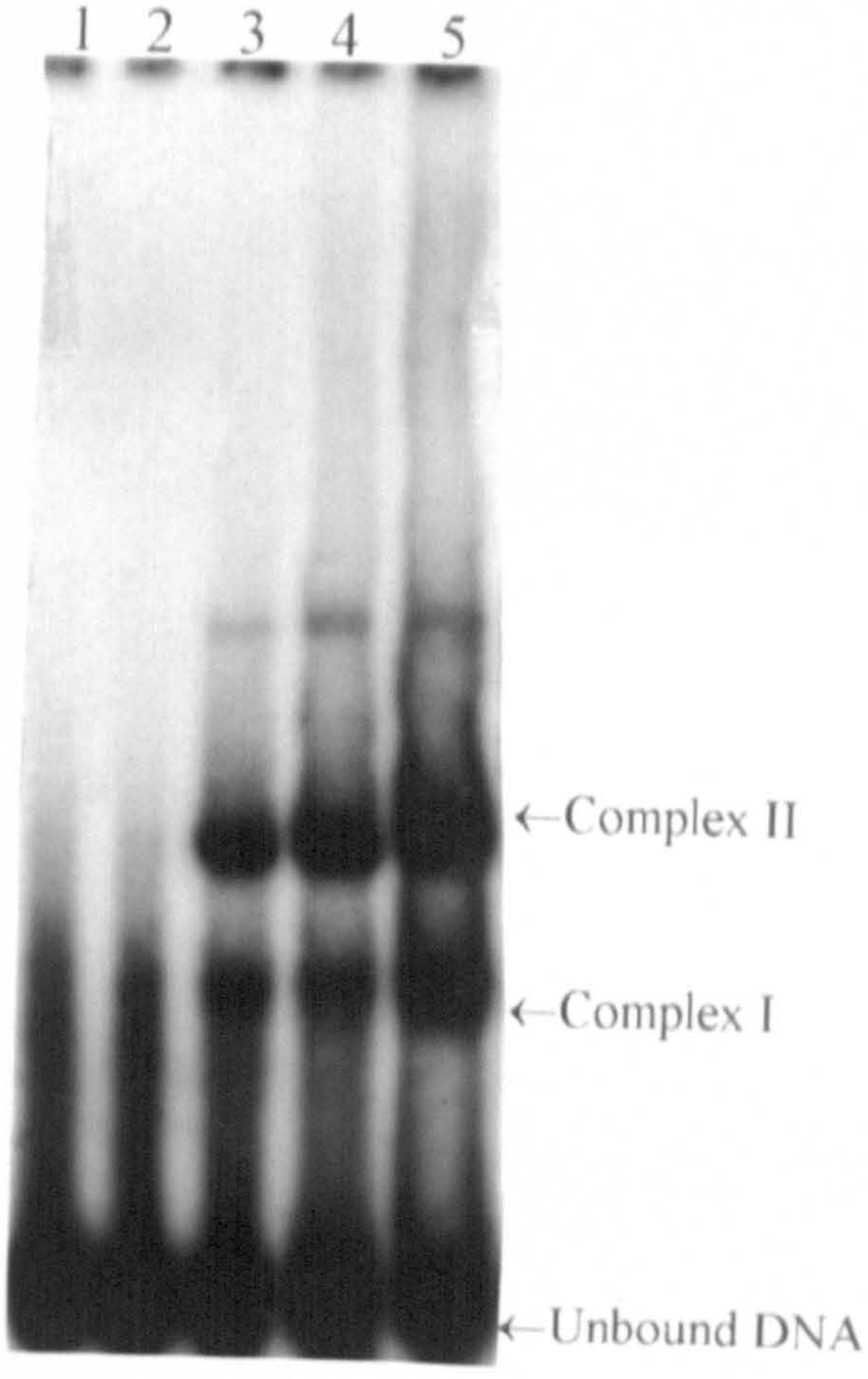
Different amounts of the recombinant Cre protein (12 ng/ml) purified from baculovirus/insect expression system were incubated with 1 ng of the radiolabelled loxP site in a band shift assay. The reactions were analysed by a non-denaturing polyacrylamide gel electrophoresis and visualised by autoradiography (Figure 4.1).

The result demonstrates that when the radiolabelled loxP site was incubated with 12 pg of the Cre protein, two distinct signals were observed (lane 3). Because Cre protein can bind either a half or a full loxP site (described in chapter 1) these signals could be Cre binding to a half loxP site (complex I) and a full loxP site (complex II) according to their different mobilities. Increasing the amount of Cre protein increased the intensities of the signals lanes 4 and 5 and the appearance of an "ill-defined" signal. No signals corresponding to either complex I or II were observed in the control reaction containing only loxP (lane 1) or when the reaction was competed with 100 fold molar excess of unlabelled loxP oligonucleotide (lane 2).

The result suggests that Cre protein can form a complex with loxP oligonucleotide under the band shift assay condition described. In previous band shift assays investigating the optimum binding conditions (results not shown), no signals were observed when 1-5 pg of the Cre protein was incubated with 1 ng of the radiolabelled loxP site. That result in addition to the observation made here suggests that the optimum amount of Cre protein required for the Cre-loxP interaction is between 5 and 12 pg. Further increase in the amount of the Cre only provides more Cre molecules to bind other loxP sites as evidenced by the increased intensity of the complexes formed.

Figure 4.1-Investigating Cre-loxP interaction in band shift assay

Different amounts of Cre protein (12 ng/ml) were incubated with 1 ng of radiolabelled oligonucleotide corresponding to the loxP site in a band shift assay. Assays were done essentially as described in materials and methods (chapter 2). Complexes formed by the recombinant Cre protein and ^{32}P-labelled loxP sites were separated by non-denaturing polyacrylamide gel electrophoresis (5%) and visualised by autoradiography. Lane1; no Cre, lane2 ; ^{32}P-labelled loxP and Cre protein (12 pg) competed with 100 fold molar excess of cold loxP oligonucleotide, lanes 3-5; ^{32}P-radiolabelled loxP incubated with 12 pg, 24 pg or 60 pg of the recombinant Cre protein respectively.


1 2 3 4 5
←Complex II
←Complex I
←Unbound DNA

The higher "ill-defined" signal observed with increasing amount of Cre protein could be the Cre-loxP complexes interacting with each other to form a Holliday structure as an intermediate for strand exchange. However, because any strand exchange would resolve back to the original complexes such a Holliday structure would be very unstable.

The band shift result observed here only shows that the recombinant Cre protein can bind DNA and gives an estimate of the amount of Cre protein required to observe protein-DNA interaction. To investigate if the recombinant Cre protein can specifically recognise and interact with oligonucleotide sequence corresponding to loxP site other DNA oligonucleotides were used as substrates for the Cre protein described in the section.

4.3.2 Investigating the specificity of the Cre protein for loxP binding site

In order to determine whether the DNA binding activity of the Cre protein is sequence-specific, band shift assays were done using radiolabelled oligonucleotide corresponding to the frt or a hybrid of loxP/frt sites as the substrate for the Cre protein.

As previously mentioned, Flp recombinase encoded by the yeast 2μ circle plasmid mediates site-specific recombination at frt sites and shows similar mechanism of site-specific recombination as the bacteriophage P1 Cre recombinase. Furthermore, the frt site like loxP is a 34 bp sequence consisting of two 13 bp inverted repeats flanking an 8 bp spacer region with each inverted repeat forming a binding site for the Flp protein (Andrews *et al.*, 1985). However, comparison of the loxP sequence and the frt sequence (shown below) does not show much conservation (identical DNA bases shown as -; degenerate codes shown by the respective alphabet [see appendix 9]). The frt sequence is therefore an interesting substrate to investigate if the structure formed by a 34 bp sequence consisting of two inverted repeats flanking an 8 bp sequence is enough for Cre binding or that Cre interaction is sequence specific.

LoxP	ATAACTTCGTATA TATTGAAGCATAT	**ATGTATGC** **TACATACG**	TATACGAAGTTAT ATATGCTTCAATA
Frt	gaagttcctatac cttcaaggatatg	**tttctaga** **aaagatct**	gaataggaacttc cttatccttgaag
Identity	RW-RY-Y-KWWWM	WWKYWWSS	K-WWM-R-RY-WY

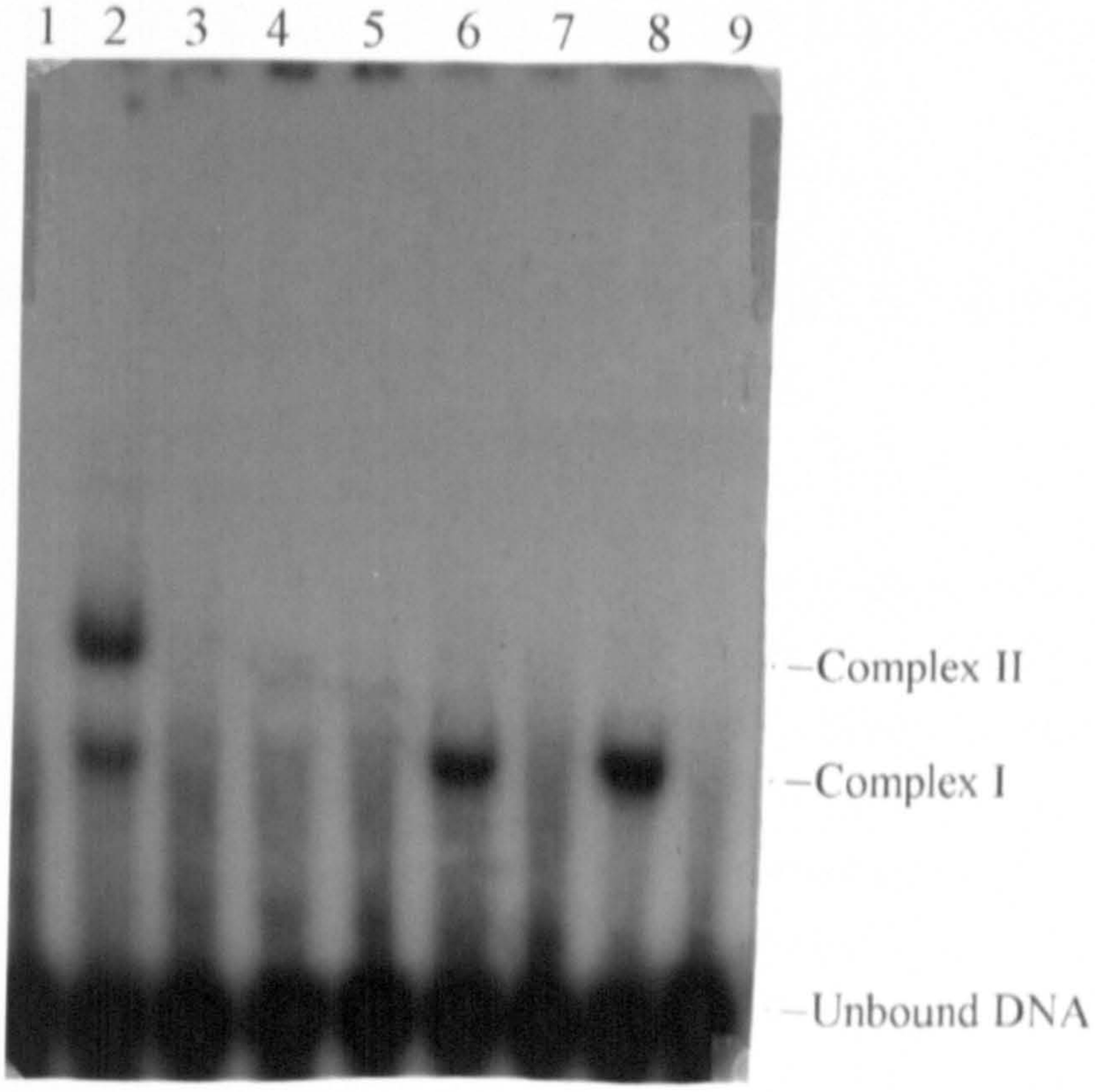
To investigate the specificity of the Cre-loxP interaction, the recombinant Cre protein (20 pg) was incubated with 2 ng of radiolabelled oligonucleotide sequences corresponding to: the full loxP site, the Flp recognition site (frt) or a hybrid of the loxP 13 bp sequence and frt 13 bp sequence flanking either the loxP spacer or the frt spacer sequence (see appendix 4 for sequences) in a band shift assay. The reactions were separated on a 5% non-denaturing polyacrylamide gel and the protein-DNA complexes visualised by autoradiography (Figure 4.2).

When Cre was incubated with radiolabelled oligonucleotide corresponding to the full loxP site, two distinct signals were observed (complexes I and II; lane 2) corresponding to Cre binding to a half loxP site (complex I) and a full loxP site (complex II) discuss below. Addition of a 100-fold molar excess of unlabelled loxP sequence efficiently competed for the Cre-loxP interaction (lane 3) as shown by the absence of any signal. When Cre was incubated with the Frt site, no signals were observed (lane 4). When the Cre protein was incubated with the loxP/Frt hybrid containing either the loxP spacer sequence or Frt spacer sequence one band was observed (complex I; lanes 6 and 8). In either case the interactions were competed for by the addition of 100-fold molar excess of the corresponding unlabelled oligonucleotide sequences (lanes 7 and 9 respectively). No signals were observed in the control reaction containing only loxP oligonucleotide (lane 1).

A number of observations can be drawn from this result. No complexes were formed between the Cre protein and the full frt site indicating that the recombinant Cre protein is very specific for the recognition loxP site. When Cre protein was incubated with loxP/frt hybrid sequence, only one signal was observed (complex I), indicating that the Cre protein bound to the loxP sequence alone. Because complex I from both half (lane 6 and 8) and full loxP site (lane 2) co-migrate, this represents Cre protein bound to half of the loxP site, whilst complex II represents Cre bound to the full loxP site confirming the initial observation that Cre forms two complexes with its recognition loxP site. Furthermore because the Cre protein formed complex I with a loxP/frt hybrid containing either the loxP spacer or the Frt spacer this indicates that the spacer sequence has no influence in the binding of Cre to the loxP site. The Cre protein in solution exists as monomers as shown by its migration as a single band of 38 kDa on non-denaturing polyacrylamide gel electrophoresis (section 3.5.4). Taking this data together, complex I could therefore represent one molecule of Cre protein bound to loxP inverted repeat, whilst complex II could represent each inverted repeat in a full loxP site bound to a single Cre molecule. This is consistent with the absolute

Figure 4.2-Investigating the specificity of the Cre protein for loxP binding site

The recombinant Cre protein (20 pg) was incubated with 10 ng of radiolabelled oligonucleotide sequence corresponding to; the full loxP site, the full Frt site, a hybrid of 13 bp loxP sequence and 13 bp Frt sequence flanking either the loxP spacer region or the Frt spacer region in a band shift assay. Complexes formed were separated by non-denaturing polyacrylamide gel electrophoresis. Band shifts were visualised by autoradiography. Lane 1 ; no Cre , lane 2; ^{32}P-labelled loxP incubated with Cre protein, lane 3; binding of Cre protein to ^{32}P-labelled loxP competed with cold loxP sequence, lane 4; ^{32}P-labelled frt incubated with Cre protein, lane 5; binding of Cre to ^{32}P-labelled frt competed with cold frt oligonucleotide, lane 6; ^{32}P-labelled loxP/frt hybrid (containing loxP spacer sequence) incubated with Cre protein, lane 7; binding of Cre to radiolabelled loxP/frt hybrid (containing loxP spacer sequence) competed with the corresponding cold oligonucleotide, lane 8; ^{32}P-labelled loxP/frt hybrid (containing frt spacer sequence) incubated with Cre protein, lane 9; binding of Cre to radiolabelled loxP/frt hybrid (containing frt spacer sequence) competed with the corresponding cold oligonucleotide sequence.


1 2 3 4 5 6 7 8 9
–Complex II
–Complex I
–Unbound DNA

stoichiometry of the Cre-loxP complex to be one molecule of Cre bound per inverted repeat, or two molecules per loxP site (Mack *et al.*, 1992).

Results from band shift analysis shows that the recombinant Cre protein does not bind to frt (Flp recombinase recognition) site under the assay conditions used. Although footprinting assays were not done to investigate this further, it is possible that Cre may form a weak interaction with the frt site. This has actually been reported to be the case. Abremski and Jayaman (unpublished observation) demonstrated that Cre binds frt sequence with a 20-fold lower affinity. If Cre protein binds frt sequence, this suggests that the frt site could be a very inefficient substrate for Cre-mediated site-specific recombination which could therefore be manipulated to achieve irreversible substrates for the Cre reaction. To investigate this, plasmid DNA substrates containing a wild-type loxP sequence and a hybrid of loxP/frt sequences were constructed, as described in the next section.

4.3.3 Construction of "mutant" substrates for Cre-mediated site-specific recombination

Two mutant plasmids, mutant A and mutant B were constructed. The mutant plasmids are identical to pAT-lacZloxP2 (section 3.6) in that they contain two site-specific recombination sequences (which are directly repeated relative to other) flanking the α peptide of *lac*Z gene. However, compared to pAT-lacZloxP2 which contained two wild-type loxP sequences, each mutant plasmid contained one wild-type loxP site and one oligonucleotide sequence consisting of a hybrid of loxP and frt half sites. In mutant A, the oligonucleotide sequence corresponds to one 13 bp sequence of loxP site and one 13 bp sequence of frt site flanking a loxP spacer whilst in mutant B, the oligonucleotide sequence corresponds to one 13 bp sequence of loxP site and one 13 bp sequence of frt site flanking an frt spacer sequence. The recombination sites are schematically represented in Figure 4.3. The mutant plasmids were designed in this way, so that if Cre protein mediates site-specific recognition between the wild-type loxP site and the hybrid loxP/frt site, the event would lead to the excision of the lacZ gene which can be analysed either directly by restriction endonuclease digestion or by phenotypic colour selection of substrates and products following an *in vivo* assay as described in sections 3.7.1 and 3.7.2.

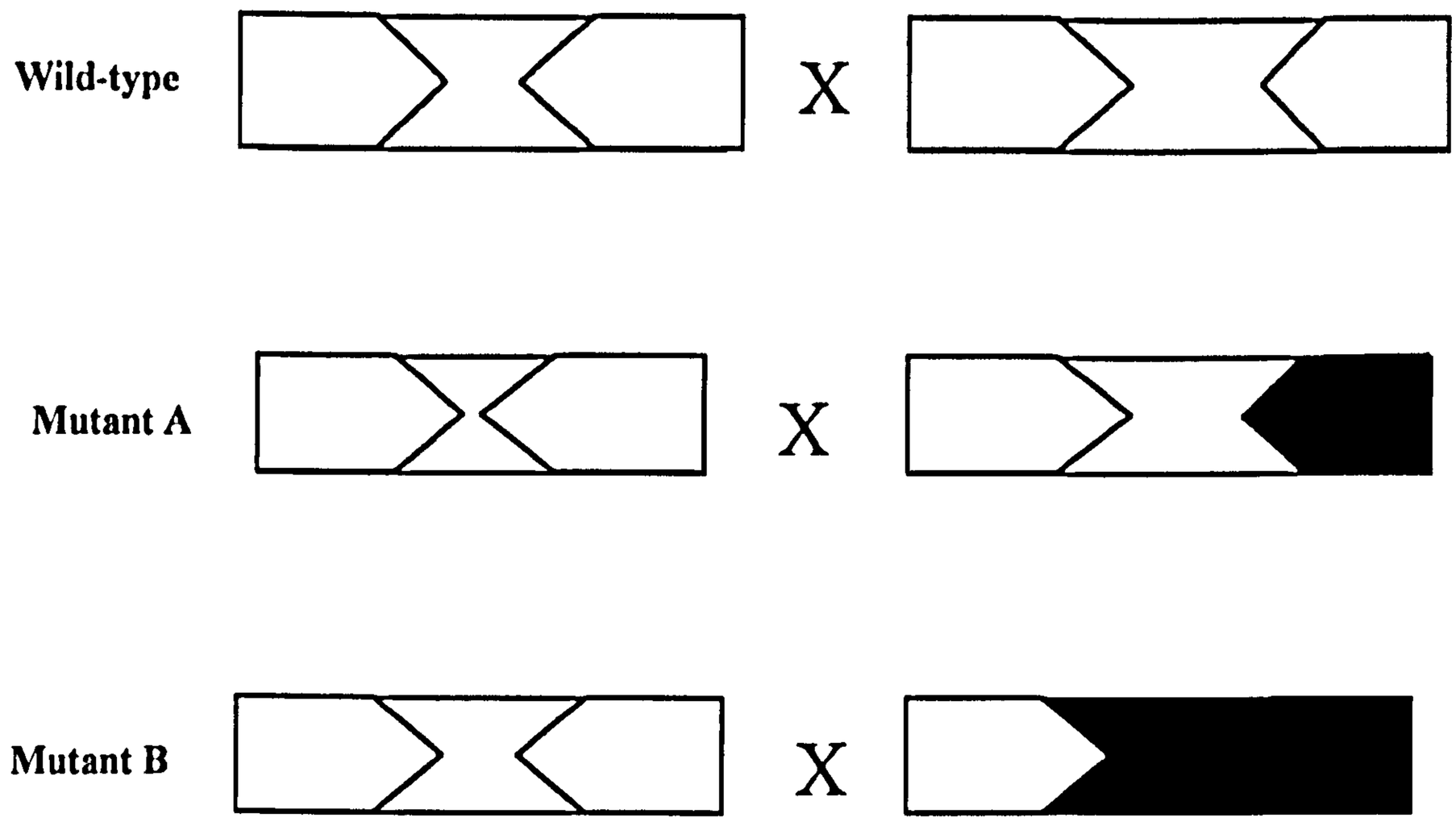


Figure 4.3- Schematic representation of intramolecular recombination sites.The wild-type recombination sites consists of two loxP sites, each made up of two inverted repeats (shown as clear boxes with arrowheads) flanking a spacer region. Mutant A recombination sitesconsists of one loxP site and a hybrid loxP site made up of 13 bp sequence of loxP (shown as clear box with arrowhead) and 13 bp sequence of frt (shown as solid box with arrowhead) flanking a loxP spacer region. Mutant B recombination sites consists of one loxP site and a hybrid loxP site made up of 13 bp sequence of loxP and 13 bp sequence of frt flanking frt spacer region.

4.3.4 Investigating Cre-mediated site-specific recombination in mutants A and B

As described above, the aim of constructing mutant plasmids was to investigate if Cre protein would recognise frt sequence as a substrate for site-specific recombination in the hope of identifying irreversible substrates for the Cre reaction.

In order to investigate this, each mutant plasmid (2 μg) was incubated with different amounts of the recombinant Cre protein in a 30 μl reaction volume for 20 minutes at 37°C. After termination of the reaction, an aliquot of each (containing about 50 ng of DNA) was transfected into *E. coli* DH5α competent cells and transformants analysed on L-agar plates (containing ampicillin, X-gal and IPTG). When mutant A was used as the substrate, all the forty-seven colonies observed on L-agar plates were blue. Also when mutant B was used as the substrate all the sixty-nine colonies observed on L-agar plates were blue. In contrast, when pAT-lacZloxP2 was used as the substrate about 47% of the colonies observed were white.

These results suggests that the efficiency of intramolecular (excision) recombination using mutant A or mutant B is less than 2% and 1% respectively. This observation in addition to band shift assays suggests that any interaction between the recombinant Cre protein and the frt site, may be too weak to detect *in vitro*. However, this may not be the only factor to explain the lack of detectable intramolecular recombination events. It is possible that in mutant A (see Figure 4.3), the Cre protein may cause strand cleavage within the two loxP spacer regions generating a Holliday intermediate, but resolution into the heteroduplex structures becomes impossible because one of the recombining partners does not have the necessary homologous sequence for strand exchange and therefore reverts back to parental. In mutant B, the Cre protein may cause cleavage in the loxP spacer region but not the frt spacer region and therefore no strand exchange can occur. However, the data presented here cannot address this. An experiment to verify the effect of two different spacer sequences would be to change one of the loxP spacer sequence in plasmid pAT-lacZloxP2 to frt spacer sequence and investigate Cre-mediated intramolecular recombination using the modified substrate. However, the failure to observe any site-specific recombination using either mutant A or B meant that the recombination sites could not be investigated further to identify irreversible substrates for the reaction.

4.3.5 "On-Off" binding assay

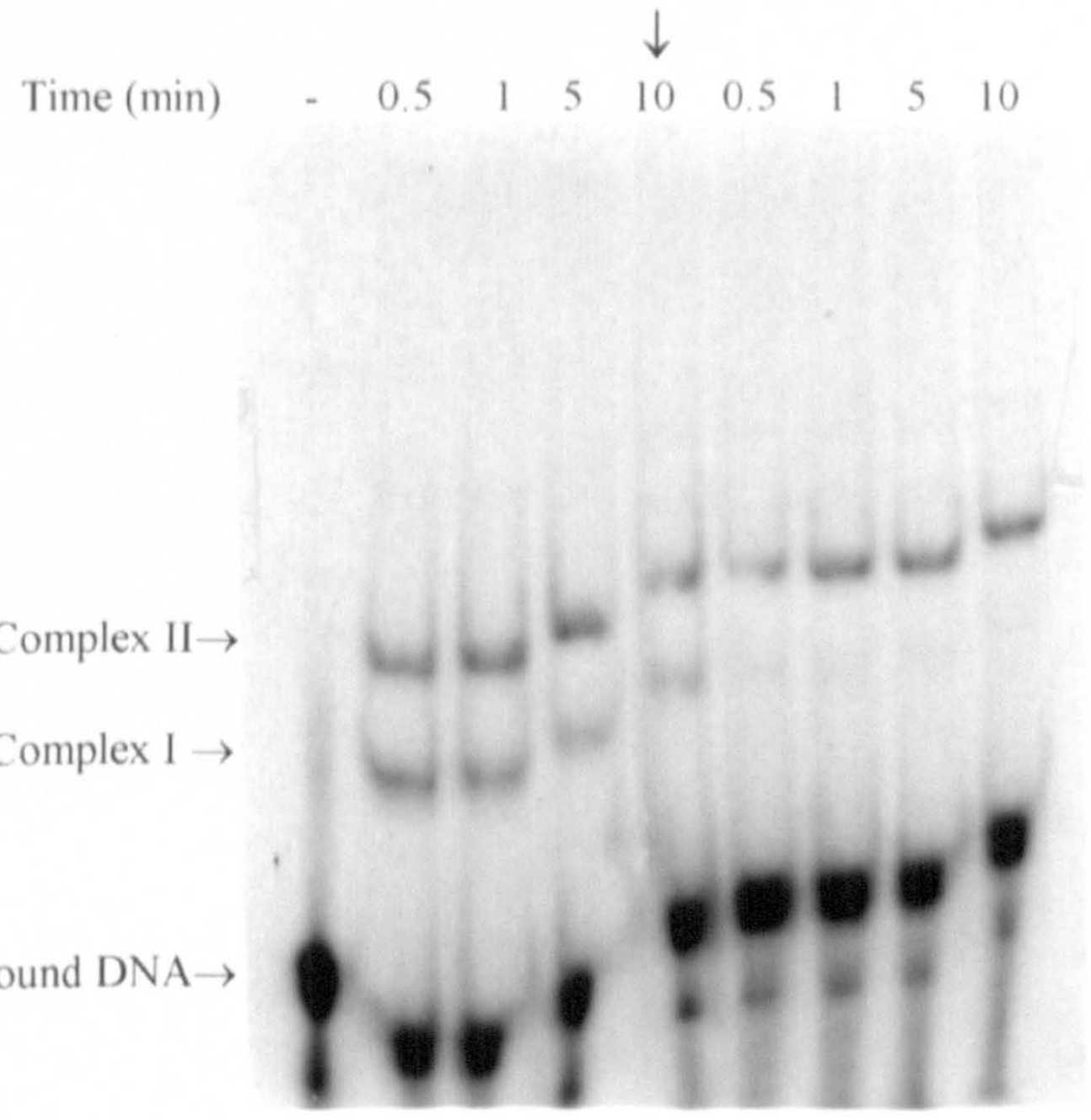
For site-specific recombination to occur, the complex formed between a single loxP site and the Cre protein must have a stable "on-rate" until it encounters another such complex for synapsis and strand exchange to take place. At the same time, the Cre-loxP complex should be able to dissociate at a given "off-rate" after the recombination event so that the subsequent heteroduplex structure formed does not revert back to the original substrates. This suggests that the stability of the initial Cre-loxP complex formed is very important if recombination is to occur. In order to investigate this, the effect of DNA competition on the "on-rate" stability of the Cre-loxP complex was analysed in a band shift assay.

The recombinant Cre protein (160 pg) was incubated with 16 ng of a ^{32}P-radiolabelled oligonucleotide sequence corresponding to a loxP site in 80 µl total reaction volume. At various time points in the reaction, 10 µl aliquots were separated on a 5% non-denaturing polyacrylamide gel. After a 10 minute incubation period, 80 ng (10-fold molar excess) of cold loxP oligonucleotide was added to the reaction. Aliquots (10 µl) were removed at various time points and immediately loaded onto the running gel. Protein-DNA complexes were visualised by autoradiography (Figure 4.4).

This result is consistent with the Cre-loxP interaction forming two complexes. These complexes were observed even when the Cre was incubated with the loxP oligonucleotide for 30 seconds. When a 10-fold molar excess of cold loxP sequence was added to the reaction (time of addition indicated by a vertical arrow), complex I was competed off within 30 seconds, whilst complex II was not displaced even after 10 minutes suggesting that complex II is more stable than complex I. Complex II represents two molecules of Cre protein bound to a full loxP site, implying a possible protein-protein interaction between the Cre molecules that stabilises the complex formed at the loxP site. The results also show that the kinetics of the Cre-loxP interaction is very fast, Cre-loxP binding occurring in less than 30 seconds, however there is no indication of how many loxP sites are occupied by Cre protein at this time point. Although the rate of dissociation of the Cre molecules from the loxP site was not determined, as it is shown in section 4.6, the *in vitro* site-specific recombination can occur in less than a minute suggesting that the "off-rate" can be as rapid as the "on-rate".

Figure 4.4-"On-Off" binding assay

The recombinant Cre protein (320 pg) was incubated with 16 ng of radiolabelled loxP site in a 80 µl total reaction volume in a band shift assay. At various time points in the reaction 10 µl aliquots were removed and separated by non-denaturing polyacrylamide gel electrophoresis. After 10 minutes incubation period, 80 ng (10 fold molar excess) of cold loxP oligonucleotide was added to the reaction (the vertical arrow indicates time of addition of excess cold loxP oligonucleotide). Aliquots (10 µl) were removed at various time points into the competition reaction and separated on the same gel. Band shifts were visualised by autoradiography.


Time (min) - 0.5 1 5 10 0.5 1 5 10
Complex II→
Complex I →
Unbound DNA→

The stability of the Cre-loxP binding interaction has important implications for the *in vivo* site-specific recombination approaches proposed in eukaryotic cells: these cells contain many DNA binding proteins which might be expected to compete with Cre for loxP sites so perhaps preventing encounters between either Cre and its loxP sites or Cre-loxP complexes. The observation that the Cre-loxP complex II is stable over at least 10 minutes suggests that the complex can be stabilised from non-specific interactions in the cell until it encounters another such complex for recombination to take place. An alternative experiment would have been to investigate the competition of total protein extract from eukaryotic cells and the recombinant Cre protein for loxP binding site in a band shift assay. Although this experiment was not done, results of Cre protein mediated targeted integration in CHO cells (described in chapter 5) demonstrates that despite possible competition from other cellular proteins for loxP binding sites, the recombinant Cre protein is able to catalyse site-specific recombination *in vivo*.

4. 4 The effect of temperature on Cre-loxP reaction

As previously described in the introduction section, there are many factors such as salt concentration and temperature that may influence protein-DNA interaction either by affecting the affinity of the protein for DNA or through conformational changes of the DNA. In order to investigate the effect of temperature on the Cre-loxP interaction, band shift assays were incubated at different temperatures.

The recombinant Cre protein (200 pg) was incubated with 10 ng of radiolabelled oligonucleotide sequence corresponding to a loxP site in a 100 µl reaction volume. Incubation was done at the following temperatures; 0°C, 10°C, 15°C, 21°C, and 37°C. At various time points in each incubation reaction, 10 µl aliquots were removed and separated on a 5% non-denaturing polyacrylamide gel and protein-DNA complexes visualised by autoradiography. The protein-DNA complexes observed at 0°C were compared with those observed at 37°C (Figure 4.5).

The band shift patterns of the Cre-loxP complex were the same for all the incubation temperatures investigated, suggesting that the binding interaction of the Cre-loxP is independent of the temperature range studied. However, Drew *et al.*, (1985) proposed that the TpA sequence in the spacer region of the loxP site may be predisposed to melting. This report, together with the observation made here, suggests that

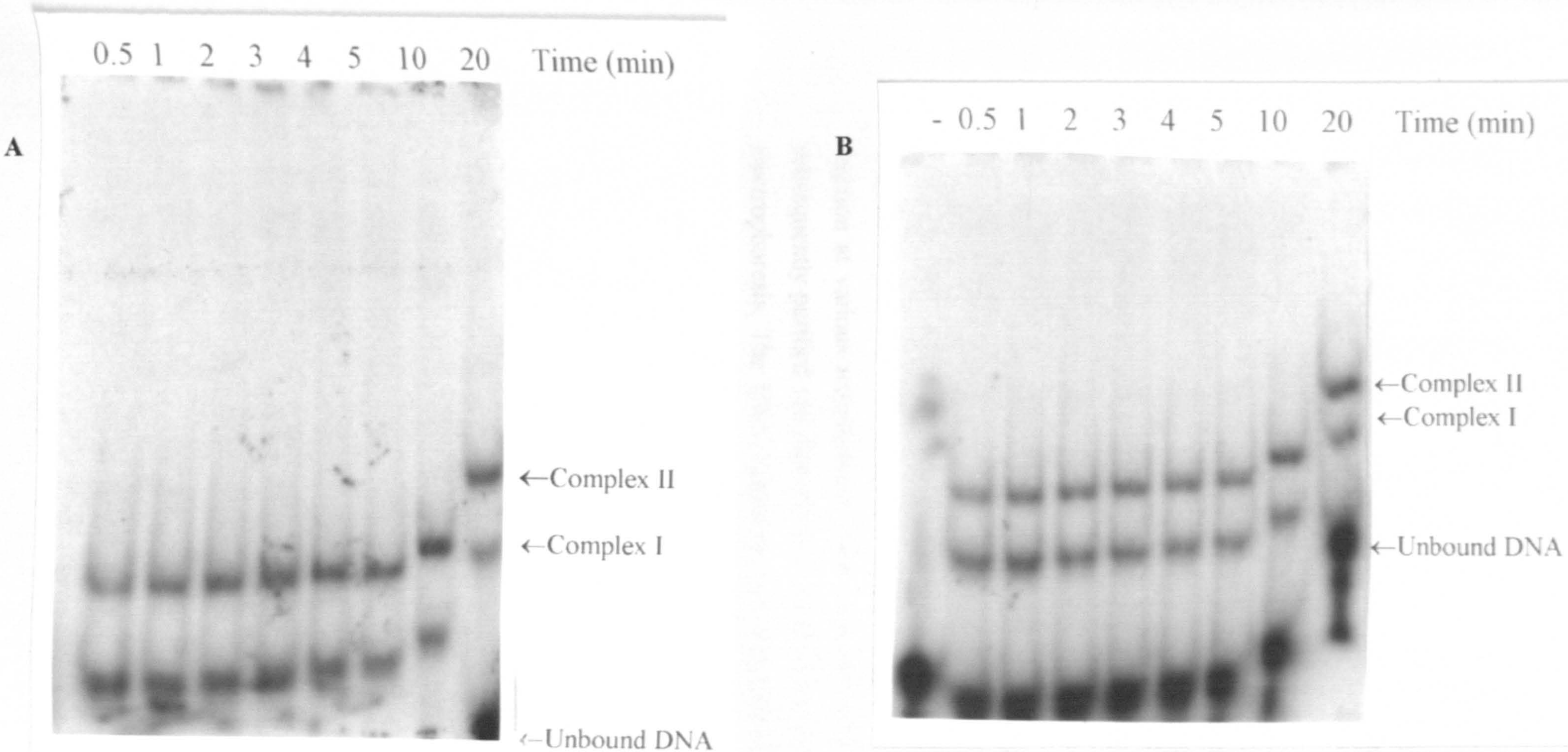


Figure 4.5- Investigating the effect of incubation temperature on the Cre-loxP interaction

The recombinant Cre protein (200 pg) was incubated with 10 ng of ^{32}P radiolabelled loxP site in a 100 µl reaction volume in a band shift assay. Incubation was done at the temperatures; 0°C, 10°C, 15°C, 21°C, and 37°C. Aliquots (10 µl) were removed from each reaction at the time points shown and separated by non-denaturing polyacrylamide gel electrophoresis (5 %). Band shifts were visualised by autoradiography. **A** corresponds to incubation at 0°C; **B** corresponds to incubation at 37°C.

temperature may be essential for other steps in the recombination process such as unwinding of the loxP site either during the pairing or the strand exchange steps. However, it is possible that the conditions chosen are not stringent enough to allow detection of temperature effect on Cre-loxP interaction. Perhaps a more accurate investigation of the Cre-loxP binding interaction at different temperatures would be to cross-link the complexes at the particular temperatures before separation on a non-denaturing polyacrylamide gel.

The observation from band shifts suggested that Cre-loxP binding is not affected by temperature, however it is possible that other steps in the recombination process may be affected by temperature. In order to investigate this, Cre-mediated intramolecular recombination assay was done at different temperatures, described in the next section.

4.4.1 Temperature effect on the efficiency of the excision reaction

The recombinant Cre protein (0.3 μg) was incubated with 2 μg of plasmid pAT-lacZloxP2 (described in section 3.6) in a complete *in vitro* recombination assay reaction at various temperatures. After termination of each reaction, the DNA was subsequently purified and digested with *Bam*HI and *Sac*I and analysed by agarose gel electrophoresis. The DNA fragments were transferred onto a nylon membrane and hybridised with a radiolabelled probe corresponding to the β-lactamase ampicillin resistance (Amp^r) gene. Southern hybridisation signals were quantified using a phosphorimager (described in section 4.6). The efficiency of recombination (% excision) which is a measure of the ratio of excised fragments to non-excised fragments is expressed as a function of temperature (Figure 4.6).

These results shows that there is a broad temperature optimum for efficient Cre-mediated excision reaction. This observed thermostability may allow the use of the recombinant Cre protein in thermophilic organisms. However, the efficiency of Cre-mediated excision decreases substantially at higher temperatures. Given that excision is a more favourable process over integration, the use of Cre-loxP system for site-specific integration in thermophilic organisms may occur at very low frequency.

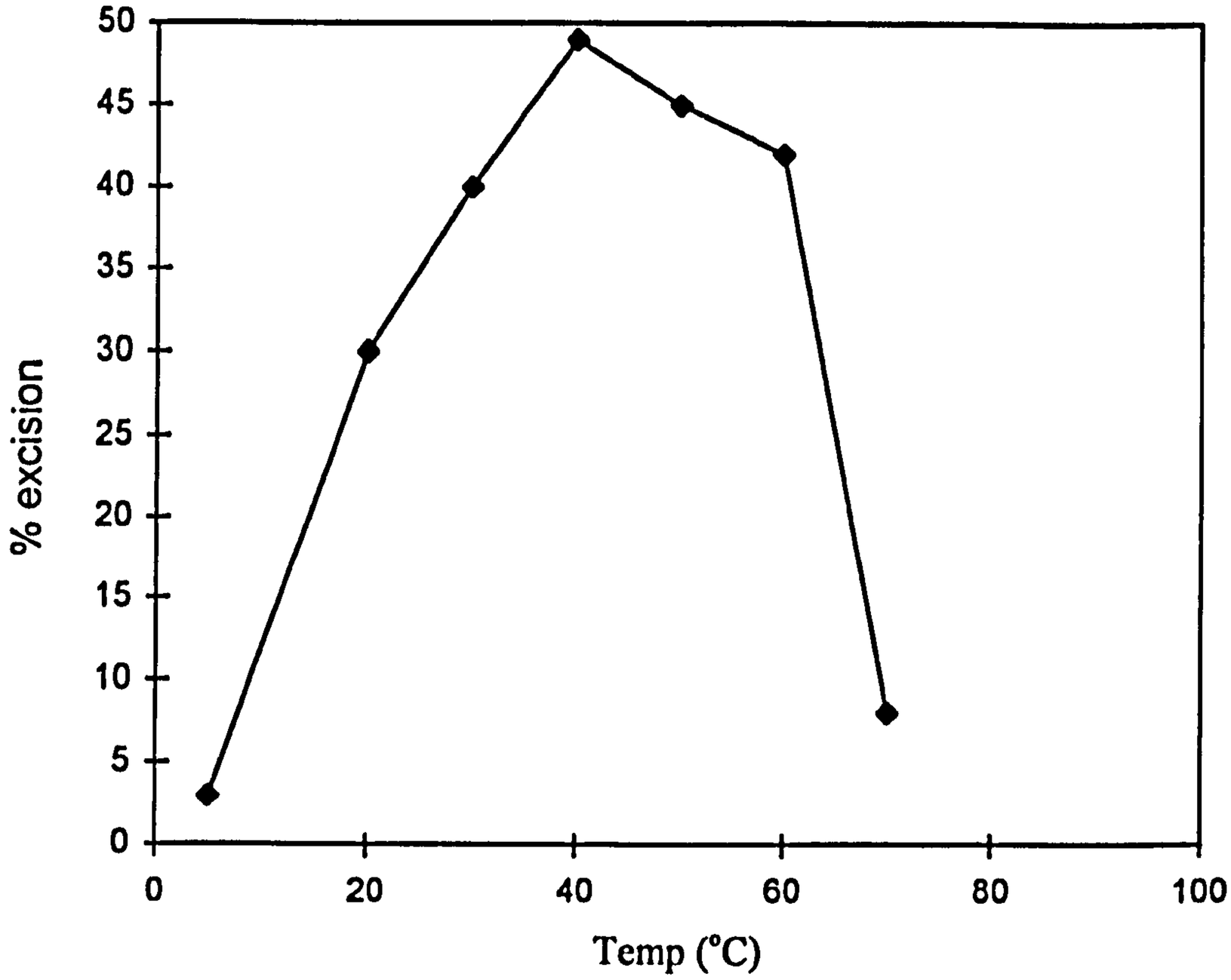


Figure 4.6- Investigating the effect of temperature on the efficiency of the excision reaction.The recombinant Cre protein (0.3 μg) was incubated with 2 μg of pAT-lacZloxP2 in a complete *in vitro* assay (see materials and methods) at various temperatures. After termination of each reaction, the DNA product was analysed by *Bam*HI and *Sca*I restriction digestion and Southern hybridisation following agarose gel electrophoresis. Hybridisation signals were quantified by a phosphoimager. Results of recombination (% excision) was displayed as a function of temperature.

4.5 Investigating other requirements for the Cre-mediated intramolecular recombination

As previously mentioned, the salt concentration in a solution can influence DNA-protein interaction and therefore recombination. In order to investigate the effect of ionic charges on the efficiency of the Cre-loxP reaction, Cre-mediated intramolecular recombination at different salt concentrations were investigated.

The recombinant Cre protein (0.3 μg) was incubated with 2 μg of plasmid pAT-lacZloxP2 in a complete assay reaction or in assay reactions having one of the co-factors omitted. After termination of each reaction, aliquots (containing ~50 ng of DNA) were transfected into *E. coli* DH5α competent cells and dilutions analysed on L-agar plates (containing ampicillin, X-gal and IPTG). The results of the recombination, observed as white and blue colonies (on L-agar plates) for recombinants and substrates respectively were quantified. The percentage recombination is expressed as a measure of the ratio of white to blue colonies (Table 4.1).

The results showed 48% optimum Cre-mediated excision reaction in the complete assay mixture. The efficiency of the reaction was not significantly affected by Na^+ ions (25 mM-150 mM), however the complete omission of Na^+ or the presence of 150 mM Na^+ significantly reduced the efficiency of recombination (the presence of Na^+ in the Cre solution might have compensated for some of the effect that would have resulted from total omission in the reaction). This indicates that Na^+ ions may be necessary for recombination by neutralising a certain portion of the phosphate groups in the loxP site and thereby increase the binding affinity of Cre. However, high concentrations of Na^+ ions may bind more phosphate groups reducing the binding affinity of Cre for the loxP site because more Na^+ ions may need to be displaced. The effect of omitting Mg^{2+} was more significant than Na^+ in the reaction. Mg^{2+} may play a similar role to Na^+ but in addition Mg^{2+} may improve the stability of the Cre-loxP binding interaction because Cre protein has been observed to dimerise in the presence of Mg^{2+} (Abremski and Hoess, 1984). Although the omission of BSA reduced recombination efficiency, the effect was not as significant as Mg^{2+}. Whilst the substitution of BSA with spermidine resulted in optimum recombination efficiency, BSA and spermidine may act differently to produce the same result. The role of BSA may act to prevent denaturing of the Cre protein as a result of non-specific interaction (e.g. with the reaction tube) whilst spermidine may act to condense the DNA substrate bringing the loxP sites into close proximity for recombination to occur.These *in vitro*

Addition	*% Recombination*
Complete	48
+ Na^+ (50 mM)	46
+ Na^+ (75 mM)	46
+ Na^+ (100 mM)	43
+ Na^+ (150 mM)	34
-Na^+	28
-Mg^{2+}	18
-BSA	28
-BSA + Spermidine	47
-Tris-HCl	46
-Cre protein	0

Table 4.1- The ionic requirement for Cre-mediated intramolecular recombination

The recombinant Cre protein (0.3 μg) was incubated with 2 μg of pAT-lacZloxP2 in an *in vitro* recombination assay. The complete reaction contained 50 mM Tris-HCl, pH 7.4; 10 mM $MgCl_2$; 30 mM NaCl; 1 mg bovine serum albumin/ml. Spermidine was added to 6 mM where indicated. After termination of each reaction, aliquots (containing ~ 50 ng of DNA) were transfected into *E. coli* DH5α competent cells and dilutions analysed on L-agar plates (containing ampicillin, X-gal and IPTG). The ratio of white and blue colonies corresponding to recombinants and substrates respectively was expressed as a measure of the percentage (%) recombination.

observations suggest that the efficiency of Cre mediated recombination is significantly affected by salt concentrations, this may not reflect *in vivo* conditions because of substantial metabolic buffering which means that the affinity of Cre for its recognition loxP sites *in vivo* may differ from *in vitro*.

4.6 Investigating the time course for the reaction

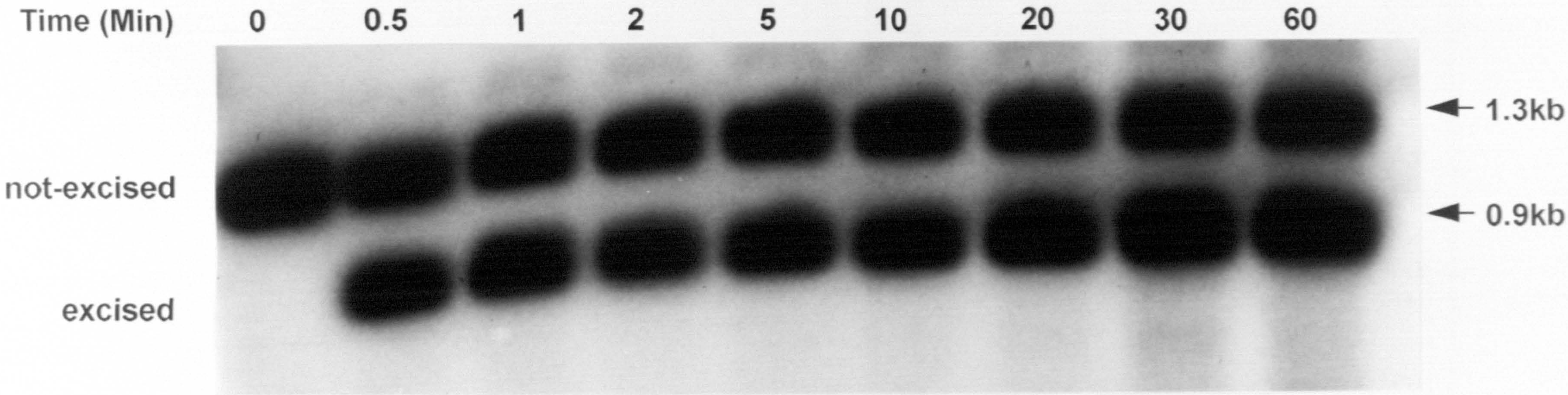
Next, the relationship between efficiency of recombination and the time course of the reaction was investigated.

The plasmid pAT-lacZloxP2 (16 µg) was incubated in a 600 µl reaction volume with 1.2 µg of the purified Cre protein. Aliquots (50 µl) of the reaction were taken at different incubation periods (0 to 60 min) and the recombination reaction terminated. The DNA was subsequently purified and digested with *Bam*HI and *Sca*I and separated on an agarose gel. The DNA fragments were subsequently transferred to a nylon membrane and hybridised with a radiolabelled probe corresponding to the β-lactamase ampicillin resistance (amp^r) gene (Figure 4.7a). As previously described in section 3.7.1., restriction endonuclease digestion of pAT-lacZloxP2 with *Bam*HI and *Sca*I yields two fragments of 2.7 and 1.3 Kbp. After Cre mediated excision of the insert between the loxP sites, restriction digestion results in the formation of a 0.9 Kbp fragment in addition to the 2.7 and 1.3 Kbp. The hybridisation signals from the non-excised (1.3 kbp) fragment and the excised (0.9 kbp) fragments were quantified with a phosphorimager. The percentage recombination (% excision) which is a measure of the ratio of excised fragments to non-excised fragments is displayed as a function of time (Figure 4.7b).

Excision products were observed as early as thirty seconds into the reaction. However, the reaction did not reach equilibrium until after five minute incubation period with only a small increase in the excision product at later time points. These results suggest that although the Cre-mediated excision reaction occurs rapidly, it takes much longer for the Cre protein to catalyse maximum excision. The observation of recombination products within 30 seconds suggests that binding and catalysis are not the rate limiting factors in recombination process, therefore another factor such as the recognition and interaction between Cre-loxP complexes could be the rate limiting step. A model reviewed in Gellert and Nash, (1987) proposes a common mechanism of bringing sites together for site-specific recombination systems that require interacting DNA

Figure 4.7a- Investigating the time course for the Cre-loxP reaction

The plasmid pAT-lacZloxP2 (16 μg) was incubated in a 600 μl volume with 1.2 μg of the purified Cre protein in a complete *in vitro* recombination assay. Aliquots (50 μl) of the reaction were taken at different incubation periods (0 min. to 60 min). The DNA was subsequently purified and digested with *Bam*HI and *Sca*I. The digested DNA products were separated on a 0.7% agarose gel, blotted onto a Nylon membrane and hybridised with a radioactively labelled probe encompassing the amp resistance gene (amp^r). The hybridisation signals were quantified with a phosphoimager.


Time (Min)
0
0.5
1
2
5
10
20
30
60
not-excised
excised
1.3kb
0.9kb

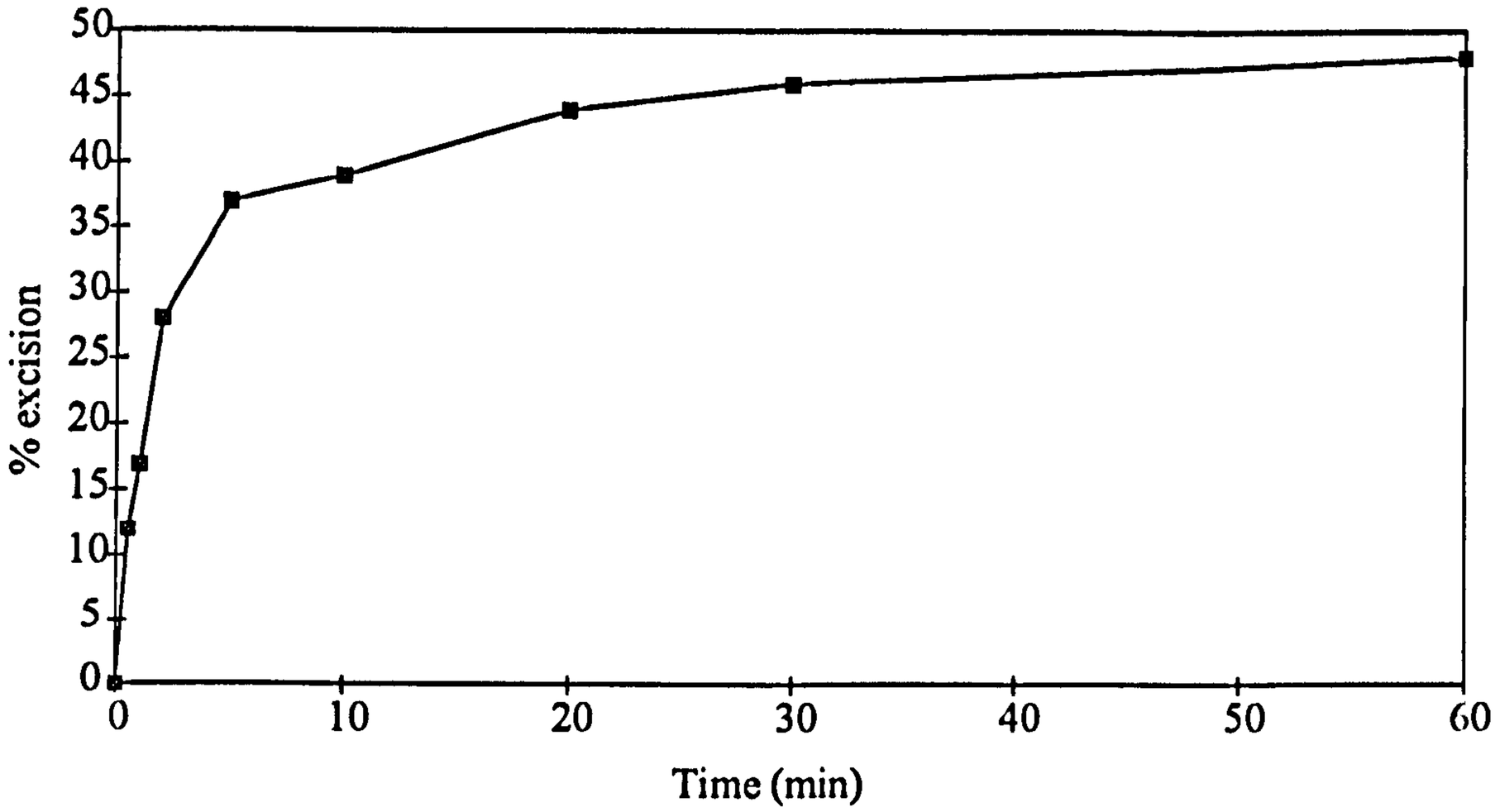


Figure 4.7b- Investigating the time course of the Cre-loxP reaction.

The percentage recombination (% excision) which is a measure of the ratio of the excised fragment to the non-excised fragment (see Figure 4.7a) is displayed as a function of time.

sequences to have specific relative orientation and site-specific recombination where relative orientation is not crucial. The model proposes that for all systems, recombination sites are juxtaposed by random collision; in the systems where orientation is not critical, all collisions lead to recombination, but in the other systems a requirement for interwrapping of the sites can place restriction on which encounters are productive. The complexity of the chromatin structure and the highly concentrated heterologous nature of the mammalian cells suggests that if the Cre-loxP site-specific recombination occurs by random collision, then such an event would occur very infrequently, however the high concentration of polycations may condense DNA structure bringing the loxP recombination sites into close proximity for recombination to take place.

4.7 Further investigation of Cre-mediated excision reaction

So far the percentage of maximum excision at equilibrium obtained using the recombinant Cre protein is 48%. This is much lower than the reported maximum excision (~70%) for native Cre protein (Hoess and Abremski, 1984) suggesting that the recombinant Cre protein may not be as active as native Cre. However, in the Hoess and Abremski study, the substrate plasmid (pRH43) used was different from that used in the present study which could explain the differences in the efficiency of excision. In order to investigate if the plasmid sequence could influence the efficiency of excision, pRH43 was used as a substrate for the recombinant Cre protein in an *in vitro* reaction.

Plasmid pRH43 (described in Abremski *et al.*, 1983) contains two loxP sites in direct repeats relative to each other. The sites are separated by approximately 2 Kbp of DNA and are flanked by two antibiotic resistance markers, amp^r and kan^r (kanamycin resistance gene; Figure 4.8a). Plasmid pRH43 (2 µg) was incubated with different amounts of the purified Cre protein in a 30 µl reaction volume for 20 minutes at 37°C. After termination of the reaction, DNA from each reaction was purified and digested with *Sca*I. The digested DNA products were separated on an agarose gel and stained with SYBR™ green solution. Restriction endonuclease digestion of pRH43 with *Sca*I yields a 4.3 Kbp fragment. After Cre mediated intramolecular recombination, restriction digestion is expected to result in the formation of a 2.0 and 2.3 Kbp fragments in addition to the 4.3 Kbp fragment. The amount of DNA in each band was quantified by fluorimager (quantification using fluorimager was found to be as

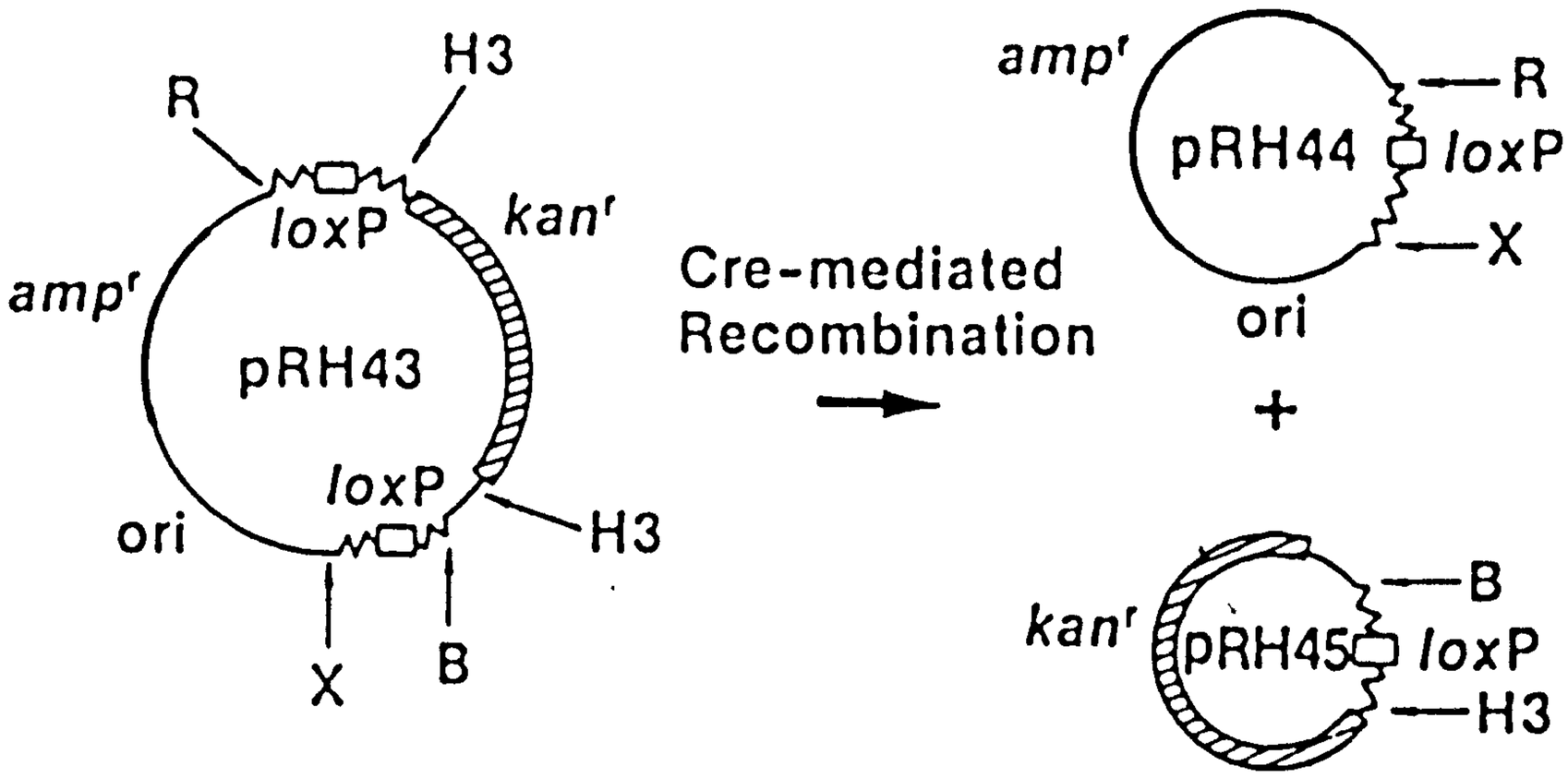


Figure 4.8a- Cre-mediated recombination in plasmid pRH43

The construction of this plasmid is described in Abremski *et al.*, (1983). Intramolecular recombination results in two smaller circular molecules, pRH44 and pRH45. Plasmid pRH44 is 2.3 Kbp in size and contains the β-lactamase gene, the pBR322 ori and a single *Eco*RI site. Plasmid pRH45 is 2.0 Kbp in size and contains the kanamycin resistance gene (kan^r) and a single *Bam*H I site. Each product contains a loxP site and hence is a substrate for Cre-mediated recombination back to the original substrate.

accurate as using phosphorimager [described in section 4.6], this method has the added advantage of directly quantifying DNA fragment in an agarose gel without Southern hybridisation). The percentage recombination (% excision) measured as a ratio of the excised (2.3 Kbp) fragment and the non-excised 4.3 Kbp fragment is displayed as a function of the amount of the Cre protein (Figure 4.8b). Because site-specific recombination excises the kanamycin resistance gene (which does not contain an origin of replication and therefore connot replicate in cells), the event can also be analysed by transfecting an aliquot of the reaction into *E. coli* DH5α competent cells and plating out transformed cells on L-agar which is supplemented with ampicillin and kanamycin. The recombination event can be assayed by the phenotypic selection of kanamycin sensitive and resistant colonies corresponding to products and substrates respectively on L-agar plates. However, this approach only gives a semi-quantitative analysis of the recombination.

A maximum excision of 63% was observed in the presence of 0.3 μg of the recombinant Cre protein. This is consistent with that achieved using purified Cre protein from *E. coli* (Hoess and Abremski, 1984). The result indicates that the recombinant Cre protein is as good as native Cre protein. The higher excision observed using pRH43 could be due to the DNA sequence in pRH43 which by virtue of its structure (for instance hairpin or kinks conformation) enhances site-specific recombination.

4.8 Intramolecular recombination between two loxP sites in inverted repeats

Results described above shows that the recombinant Cre protein can mediate site-specific recombination events between two loxP sites directly oriented to excise the intervening sequences. In order to investigate if the recombinant Cre protein can also mediate inversion, plasmid pAT-lacZloxP2$_{inv}$ was used as a substrate for the Cre-mediated recombination event.

Plasmid pAT-lacZloxP2$_{inv}$ is the same as pAT-lacZloxP2 except that the loxP sites flanking the lacZ gene are inverted relative to each other. Plasmid pAT-lacZloxP2$_{inv}$ (2 μg) was incubated with different amounts of the Cre protein in complete assay reaction. After termination of the reaction, the recombination products were analysed in two ways:

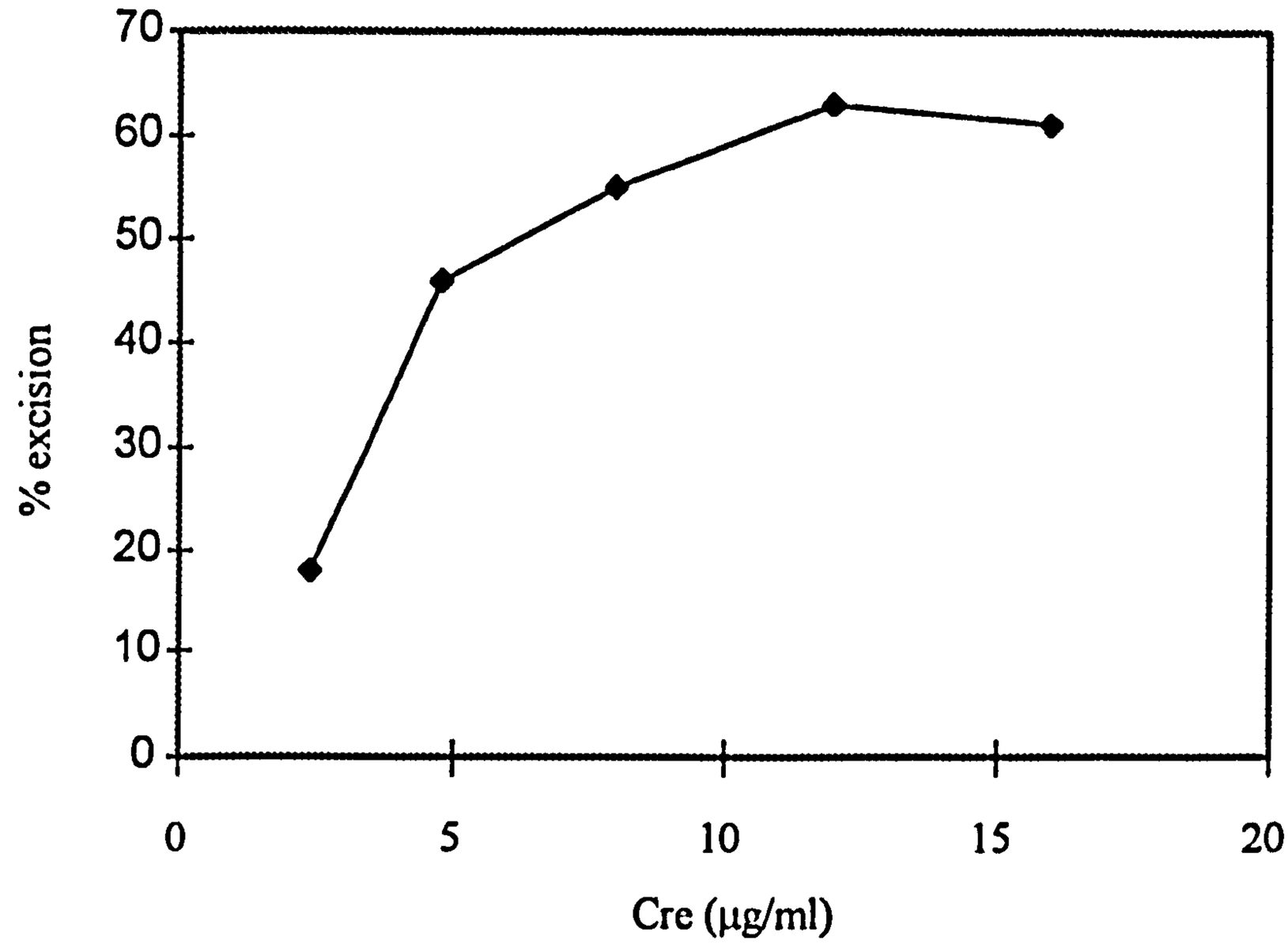


Figure 4.8b- Investigating the efficiency of the Cre-mediated intramolecular excision event.Plasmid pRH43 (2 μg) was incubated with different amounts of the recombinant Cre protein in a complete in vitro assay reaction. After termination of the reaction, DNA was purified from each reaction and digested with *Sca*I. The digested DNA products were separated on a 0.7% agarose gel and stained with SYBR^{TM} green solution as described in materials and methods. The amount of DNA in the excised (2.3 Kbp) fragment and the non-excised (4.3 Kbp) fragment was quantified using a Flourimager.

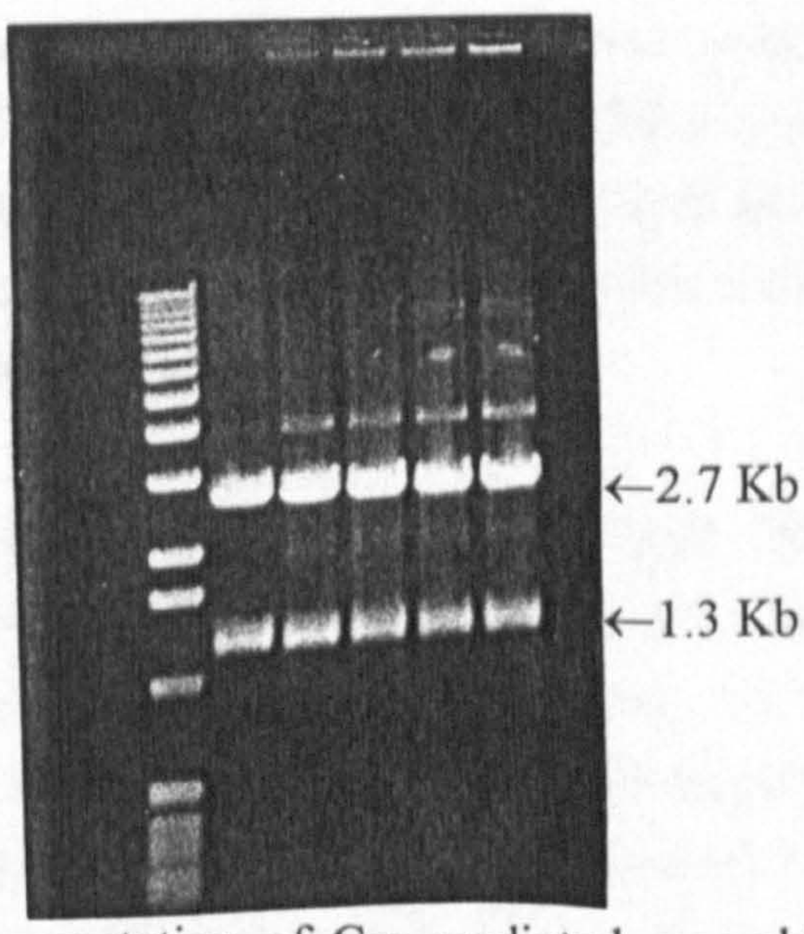
(i) DNA purified from half of the reaction volume was digested with *Bam*HI and *Sca*I and separated on an agarose gel (Figure 4.9a). Restriction endonuclease digestion of pAT-lacZloxP2$_{inv}$ with *Bam*HI and *Sca*I yields two fragments of 2.7 and 1.3 Kbp. If Cre-mediated recombination excises the insert between the two loxP sites, restriction digestion is expected to result in the formation of a 0.9 Kbp fragment in addition to the 2.7 and 1.3 Kbp. However none of the 0.9 Kbp fragment was seen suggesting that the recombination event does not lead to excision of the lacZ gene.

(ii) To confirm this observation, an aliquot of the reaction (containing about 50 ng) was transfected into *E. coli* DH5α and plated onto L-agar (containing ampicillin, X-gal and IPTG). If Cre mediated recombination excises the lacZ gene between the loxP sites, the recombination event should give rise to the phenotypic white and blue colour selection for products and substrates respectively (as described in section 3.7). Only blue colonies were observed in all the Cre mediated reactions supporting the initial observation that recombination between loxP sites inverted relatively to each other does not excise the intervening DNA fragment. However, the control reaction (absence of Cre protein) also gave blue colonies, therefore to verify that the blue colonies could also result from Cre-mediated inversion, plasmid DNA prepared from five blue colonies were sequenced using an oligonucleotide sequence designed to prime within the β-lactamase (amp^r) gene in the vector backbone. Results of the sequencing analysis (data not shown) showed that one out of the five DNA preparations had the lacZ gene inverted between its flanking loxP sites (Figure 4.9b). A more readily assayable experiment to investigate the inversion event would have been to construct a plasmid containing the loxP sites in inverted repeat flanking on the one side the lac promoter and on the other side the *lac*Z gene. The plasmid would have been designed such that the lac promoter and the *lac*Z gene are in opposite orientation with respect to each other. Cre mediated recombination at the loxP sites should invert the lac promoter between the loxP sites so that the recombination event leads to the expression of the *lac*Z gene under the transcriptional control of the lac promoter. The recombination event could be quantified by the phenotypic colour selection of substrates and products, which in this case would give rise to white and blue respectively.

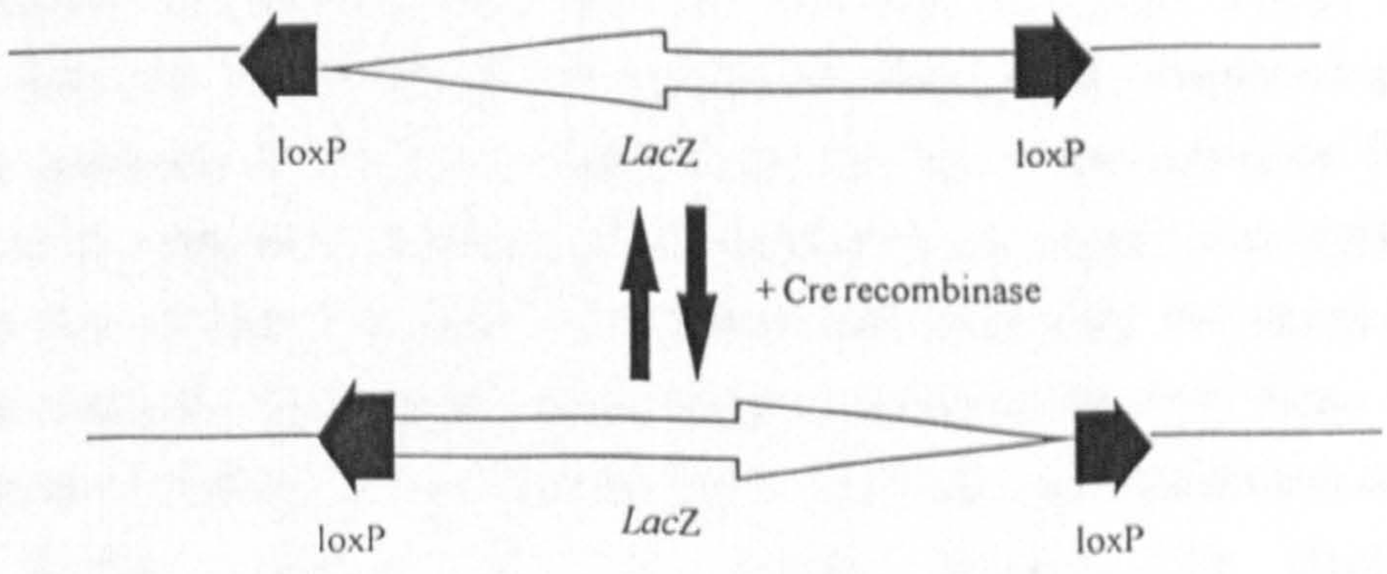
Nonetheless, this result shows that the recombinant Cre protein mediated recombination between loxP sites inverted relative to each other leads to the inversion of intervening DNA sequences at a frequency of 20%, although no efforts were made to optimise this.

Figure 4.9- Investigating Cre mediated recombination between loxP sites in inverted repeats.

A-Restriction digestion of recombination event between loxP sites in inverted repeats.The plasmid pAT-lacZloxP2$_{inv}$ was incubated with various amounts of the purified Cre protein in a 30 µl reaction volume containing 50 mM Tris-HCl, pH 7.4; 10 mM $MgCl_2$; 30 mM NaCl; 1 mg/ml bovine serum albumin for 20 minutes at 37°C. The reactions were then purified by phenol/chloroform extraction followed by ethanol precipitation. The plasmid DNA were subsequently digested with *Bam*HI and *Sca*I and separated on 0.7% agarose gel. DNA was visualised by ethidium bromide staining. M is 1 Kb DNA standard marker.



B-Schematic representation of Cre mediated recombination between two loxP sites inverted relative to each other. The Cre mediated recombination leads to the inversion of sequence between the loxP sites.



4.9 *In vitro* analysis of integration event

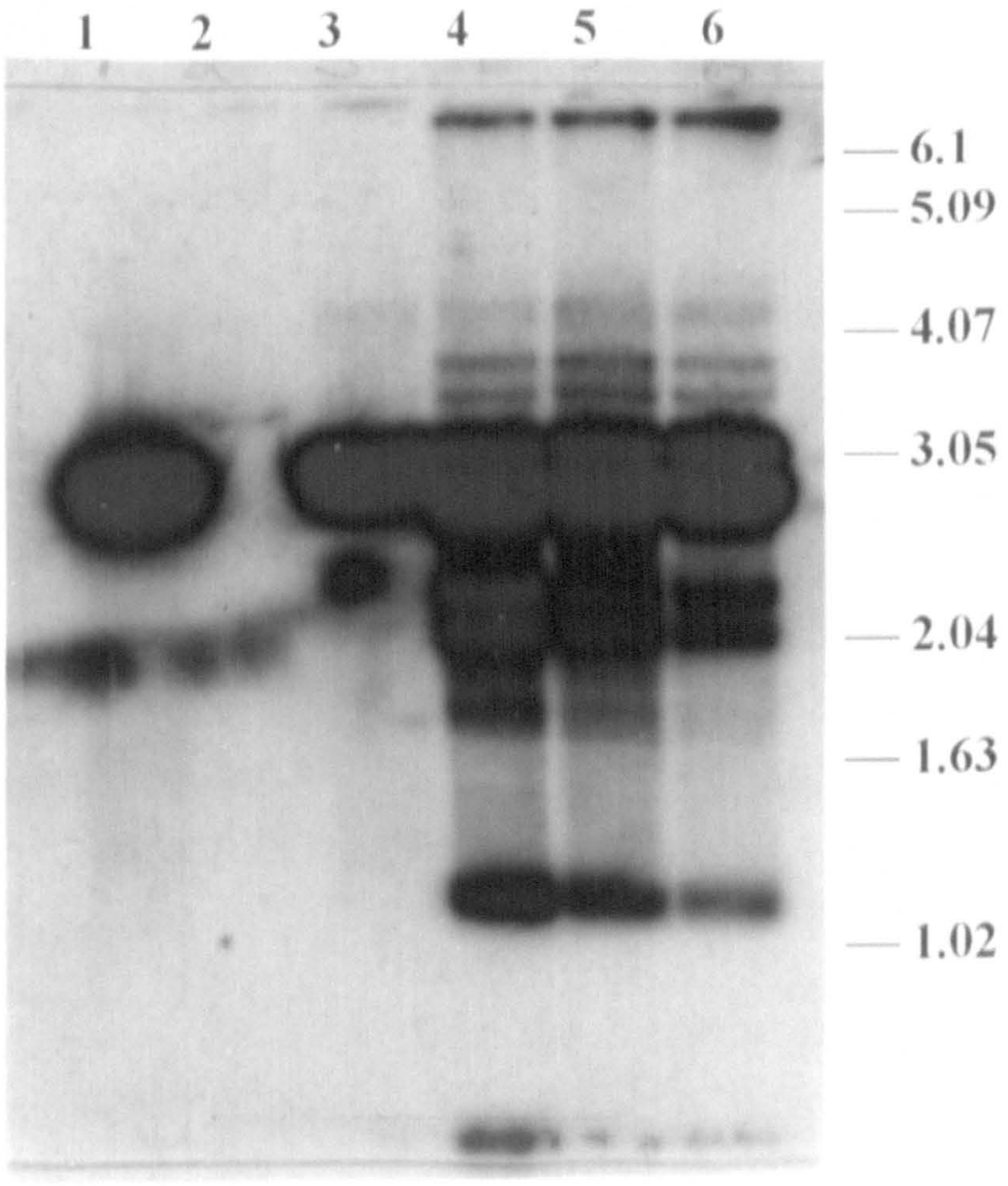
As previously mentioned, one of the objectives of this work was to directly introduce the recombinant Cre protein into cells to mediate targeted integrations (described in chapter 5). It was therefore important to initially establish that the recombinant Cre protein could mediate site-specific integration *in vitro*.

Two plasmids (described in Peakman *et al.*, 1992) each containing a single loxP site were used to investigate Cre mediated integration event. Plasmid pKSloxP was constructed by cloning an oligonucleotide pair corresponding to the sequence of loxP into the *Kpn*I site of pKS (Bluescript +, Stratagene). Plasmid p36Clox was constructed by cloning an oligonucleotide pair corresponding to the sequence of loxP between the *Eco*RV and *Kpn*I sites of the baculovirus transfer vector p36C (Page, 1989). Plasmids pKSlox and p36Clox were digested with *Sal*I and *Pvu*II respectively. The linear DNAs (2 μg each) were mixed in the absence and presence of varying amounts of the recombinant Cre protein and incubated for 20 minutes at 37°C. After termination of the reaction, the products were electrophoresed on an agarose gel and analysed by Southern blotting using a 445 bp *Pvu*II probe corresponding to the lox containing fragment of pKSlox (Figure 4.10).

Cre-mediated site-specific integration between the loxP sites of the linearised substrates, pKSlox and p36lox is expected to give rise to product species of molecular sizes; ~ 0.5 kbp, ~ 1 7 kbp, ~ 2 kbp, ~ 3 kbp, ~ 4 kbp and higher molecular weight sizes if multiple integration occurred. The results show a signal of 3.05 kbp in lane 1 (positive control) corresponding to the 445 bp probe hybridising to the complementary sequence in the pKSlox fragment. No signals were observed in lane 2 (negative control) because the p36Clox fragment does not contain the 445 fragment sequence. When pKSlox and p36Clox were incubated together in the absence of Cre protein (lane 3), only the 3.05 kbp signal was observed suggesting that no site-specific recombination had occurred between the loxP sites. In the presence of Cre recombinase (lanes 4-6), hybridisation signals of ~1.1 kbp, ~1.7 kbp, ~2 kbp, 3.05 kbp, ~4 kbp were observed. These signals corresponds to site-specific integration products. However, the presence of 3.05 kbp suggests that not all of the substrate fragments were converted to integration products. The intensity of the signals was highest when 0.3 μg of the Cre protein was used. This shows that increasing the amount of Cre protein in the reaction increases the recombination products because there are more Cre molecules to recombine more substrate DNA. However, an equilibrium is reached when the further addition of Cre protein favours the excision of

Figure 4.10- Investigating Cre-mediated integration event

2 μg of *Sal*I digested pKSlox and 2 μg of *Pvu*II digested p36clox were mixed in the presence and absence of varying amounts of the recombinant Cre protein in a complete *in vitro* assay. After termination of the reaction, the products were electrophoresed on an agarose gel and analysed by Southern blotting using a 445 bp *Pvu*II probe corresponding to the lox containing fragment of pKSlox. Lane 1; pKSlox, lane 2; p36clox, lane 3; pKSlox + p36clox, lane 4-6; pKSlox + p36clox in the presence of 0.2, 0.3 or 0.5 μg of the recombinant Cre protein respectively. The position of molecular weight markers (in Kbp) are shown on the left.


1
2
3
4
5
6
6.1
5.09
4.07
3.05
2.04
1.63
1.02

the integrated molecules back to the original substrate as shown by the weaker signals and the disappearance of some signals (~1.7 kbp) in the reaction containing 0.4 μg of Cre protein.

It is clear from the results that the recombination is dependent on both the presence and amount of Cre and also that the loxP sequence in this context is an efficient substrate for the Cre protein. The observation that maximum integration was detected in the presence of 0.3 μg of Cre protein correlates with the maximum excision (see section 3.8) for the same amount of substrate DNA. However the efficiency of the integration event was not determined to compare with the excision event.

4.10 Summary

In this chapter, the baculovirus/insect derived Cre protein was used in two *in vitro* assay systems to investigate the kinetics of the Cre-loxP site-specific recombination reaction: First, band shift assays showed that the recombinant Cre protein is very specific for its recognition loxP sites with the stoichiometry of the interaction being two Cre molecules per loxP site. Band shift assays and Cre-mediated intramolecular recombination (discussed later) suggests that the Cre-loxP binding and dissociation must occur very quickly because recombination products were detected in less than thirty seconds. However, using oligonucleotide competitors in band shift assays, results suggests that a loxP site bound by two molecules of Cre was stable for over ten minutes indicating that even though the recombination occurs very fast, a Cre-loxP complex is stable until another complex is recognised for recombination to take place.

Second, the recombinant Cre protein mediated intramolecular (excision/inversion) and intermolecular (integration) recombination in DNA substrates containing loxP sites. In the context of the substrates described in this chapter, recombination was absolutely dependent on the presence of the Cre protein and two functional loxP sites. For the intramolecular event, the recombination products were dependent on the oricintation of the loxP sites relative to each other. Cre mediated recombination between loxP sites in direct repeats led to excision, whereas the recombination between loxP sites in inverted repeats gave rise to the inversion of the intervening DNA sequence. The *in vitro* reaction was dependent on the concentration of salts, especially Mg^{2+}. Perhaps Mg^{2+} effects recombination by increasing the affinity of Cre for the loxP site and/or stabilising the Cre-loxP binding interaction. However, Cre-loxP binding was not

affected by temperature and furthermore, optimum recombination occurred over a wide range of temperatures suggesting that the Cre-loxP reaction is thermostable.

Because Cre-loxP site-specific recombination is reversible, and excision is kinetically favoured over bi-molecular integration, integrated products are inherently unstable in the presence of the recombinase. This would be a major drawback for site-specific DNA integration. An attempt to improve the stability of integrated DNA molecules was to introduce selective base changes into recombination sites such that the product sites of a recombination reaction are less likely to undergo further recombination (Albert *et al.*, 1995). In this chapter, substrate plasmids containing a wild-type loxP and a hybrid loxP/frt site were constructed to investigate if the recombinant Cre protein would mediate intramolecular recombination between these sites, in order to manipulate these hybrid sites further to improve the integration event. Unfortunately no detectable recombination events were observed between the wild-type and the hybrid loxP/frt sites suggesting that the loxP/frt hybrid was not a substrate for the recombinant Cre protein. A drawback in manipulating the loxP site to improve the stability of integrated DNA, is that it may reduce the frequency of intermolecular recombination because the integration event is less efficient than excision, therefore any enhanced stability that will be conferred upon integrated molecules by manipulating the loxP sites will probably be outweighed by even fewer integration events.

The results from the *in vitro* studies shows that the recombinant Cre protein is as efficient as native Cre in mediating site specific recombination at loxP sites, and suggests that the six histidine residues and the nuclear localisation signal of the fusion protein do not interfere with the recombinant Cre activity. This raises the potential that the histidine domain could be used as an independent molecular tag for *in vivo* analysis of Cre mediated recombination whilst the nuclear localisation signal should increase the rate of nuclear transport of the Cre protein in mammalian cells for targeted integrations, described in chapter 5.

CHAPTER 5

Targeting of exogenous DNA into CHO chromosome mediated by the Cre-loxP system

5.1 Introduction

As previously described in chapter 1, the main problem involved in the use of mammalian expression systems as host for the high level expression of recombinant proteins is that the expression levels of recombinant proteins show extreme variability between individual clones isolated from a single transfection. This is due to a differing number of copies integrated into the genome and also from chromosomal position effects. As a result an extensive number of clones must be screened in order to isolate a high-expressing clone.

A strategy that was initially described by Fukushige and Sauer (1992) used the Cre-loxP site-specific recombination system to eliminate both the constraints of chromosomal position effects and copy number variation on gene expression. In this strategy a loxP site was introduced into a region in the chromosomal such that transgenes could be targeted to that region, a process catalysed by the Cre recombinase. Independent clones isolated from a single transfection all expressed similar amount of recombinant protein showing that the Cre-mediated site-specific integration can facilitate the construction of isogenic cell lines and thereby permit reproducible gene expression in stably transformed cell lines.

Such a strategy would be advantageous if the site to which the foreign gene is to be targeted is favourable for high expression leading to the isolation of transfected cell lines expressing relatively constant, high levels of recombinant protein. This chapter investigates the potential of the Cre-loxP system for the construction of a cell line that could be used for the reproducible high level expression of recombinant proteins. Because this chapter investigates mammalian cells for recombinant protein expression, the various components involved are described in this section.

5.1.1 Choice of cell line

The wide variety of transcriptional control elements and selectable markers now available permit the use of essentially any cell line for the purposes of heterologous gene expression. However, the consideration for the preferred cell line, depends on a number of factors which includes the availability of appropriate expression vectors (described below), growth properties such as doubling time and serum requirements, stability and expression levels. Furthermore, tissue-specific or temporally regulated gene expression may require appropriate cell lines which express tissue-specific cellular factors and genetic regulatory elements to produce normal patterns of expression. For example, immunoglobulin (Ig) genes are usually expressed in lymphoid cells were they are transcribed and translated into biologically active proteins (Dorai and Moore, 1987).

5.1.2 Methods for introducing transgenes into mammalian cells

As previously described in chapter 1, there are two general methods of introducing transgenes into mammalian cells; those mediated by virus infection and those mediated by direct DNA transfer (reviewed in Kaufman, 1990). One advantage of the former relates to its efficiency; it is not unusual to achieve close to 100% DNA delivery to cells by exploiting the ability of the virus to efficiently infect cells. A disadvantage of this approach, and one that limits its utility, lies in the relative difficulty of its application; it requires much more effort to engineer a virus to express a foreign gene in a manner that does not compromise its function than it does to administer DNA directly. Also the use of viral vectors can raise a number of regulatory issues such as containment of genetically engineered viruses. As such, most attempts to express heterologous genes in mammalian cells rely on the construction of DNA vectors and their direct transfer by a number of chemical and physical methods, each with its own advantages and disadvantages (reviewed in Kriegler, 1990). Ray and Cage (1992) showed that, at least for the two cell lines and three transfection methods they studied (i.e. established Rat-1 cell line and rat primary skin fibroblasts; by $CaPO_4$ co-precipitation, lipofection and electroporation) the transfection method does not affect the level of foreign gene expression achieved.

Direct transfer methods have been used successfully in a wide variety of cell types, including Chinese Hamster Ovary cell line, CHO-DUKX B11 (Page and Sydenham, 1991) and Baby Hamster Kidney cell line, BHK-21 (Hippenmeyer and Highkin, 1993).

The efficiencies of these techniques, whilst not as high as with viral systems can approach ~90% of the transfected cells, which will continue to express the protein of interest for a period of several days and perhaps a few weeks. The transfected DNA is eventually lost unless a procedure is used to select for cells which have integrated the DNA into their genome. The limiting factor is not the uptake of the DNA into the cells, but rather the integration of the transfected DNA into the chromosome. Therefore a selection procedure is required in order to select for stably transfected cell lines. For this reason, expression vectors (described in the next section) are often designed to contain the transgene and a selectable marker within the same plasmid.

5.1.3 Design of expression vectors

The basic elements for the control of heterologous gene expression consists of a promoter-enhancer element which is responsible for the transcription of the cistron, and a polyadenylation signal which directs the cleavage and polyadenylation of the transcript (Kriegler, 1990). Many vectors also contain one or more small introns, in the belief that intron-containing vectors are more efficient for gene expression. Following the promoter element is the coding sequence of the gene or cDNA to be expressed. Mammalian cells do not show stringent requirement for translational initiation, although they do exhibit a preference for the first translational initiation codon, AUG at the 5' end of the message and especially for initiator codons having the consensus sequence .....CCACCAUGG.... (Kozak, 1989). Downstream AUGs can be recognised in mammalian cells particularly if the upstream AUG is followed by a termination codon or if the downstream AUG is preceded by IRES as found in encephalomyocarditis virus (EMCV; Kaufman, 1990). Following the gene insert is a sequence which allows for the cleavage and addition of the polyadenylation tail to the transcript. The polyadenylation signal consists of the hexanucleotide AAUAAA located between 20 and 30 nucleotides upstream of the site of polyadenylation addition (Wickens, 1990). Lack of polyadenylation signal can severely reduce the amount of cytoplasmic mRNA and translated protein. Most commercial vectors are designed to contain multiple cloning sites, a phage origin of replication (ori) sequence and multiple drug-resistance markers.

In addition to the heterologous gene, some type of selectable marker must be added on a second cistron to allow for the selection of stable cell lines. The selectable marker can be either on the expression plasmid or contained on a second plasmid. For transient expression systems or high-efficiency gene-transfer systems (i.e. viral vectors or

microinjection) a selectable marker is not needed. The first selectable markers used were based upon complementation systems encoding the hypoxanthine phosphoribosyl transferase (*hgprt*) gene and the herpes thymidine kinase (*tk*) gene (Wigler *et al.*, 1977; Roberts *et al.*, 1982). Somatic mutants lacking the *tk* or *hgprt* gene could be easily selected for by culturing the cells in the presence of nucleoside analogues such as 8-azaadenine or 6-thioguanine so that metabolism of the analogue results in the production of toxic nucleotides. The development of CHO mutants lacking dihydrofolate reductase (DHFR) allowed the use of SV40-DFHR chimaeric plasmids to select for cells able to grow in the absence of thymidine, glycine and hypoxanthine (Urlaub, 1980). The ease with which the expression levels of these CHO cells could be further increased by methotrexate selection has been the primary reason why CHO-DHFR system has become the most popular mammalian system used in the production of recombinant proteins (described in detailed below).

The main limitation of vectors containing such selectable markers is that the host cells are limited to mutant cell lines. Although TK and HGPRT mutants can be easily selected for and against, $DHFR^-$ cell lines are very difficult to obtain. For these reasons, dominant selectable markers, able to be used in any cell type, have been derived from bacterial genes which have no equivalent sequences in mammalian cells. The most widely used dominant marker is derived from the transposon encoded aminoglycoside 3' phosphotransferase (APH), which confers neomycin (neo) resistance (Jimenez and Davies, 1980). The related aminoglycoside G418 (geneticin) is extremely toxic to mammalian cells and serves to inhibit protein synthesis. Cells expressing even low levels of the neomycin gene are able to survive in the presence of high levels (mg/ml) of G418. The selection works well on a wide variety of cell types, including virtually all of the cells used for recombinant DNA (rDNA) expression. Almost as widely used is the gene encoding hygromycin B phosphotransferase, an *E. coli* gene which acts similarly to the neomycin gene (Colbere-Gerapin *et al.*, 1987). It confers resistance to hygromycin, an aminoglycoside antibiotic which affects protein synthesis by blocking ribosome translocation. Dominant selectable markers based on mammalian genes have also been developed. The three most widely used systems utilises DHFR (described later), adenosine deaminase (Yeung *et al.*, 1983) or glutamine synthetase (Young and Ringold, 1983). All these components were taken into consideration in constructing a cell line that could be used for the reproducible high level expression of recombinant proteins.

In this study, CHO cells (Levinson, 1990) were used as the host system for the expression work, because CHO cells are well suited for the generation of production

cell lines for many reasons. Chief among these is the availability of CHO cells deficient in their ability to produce dihydrofolate reductase (DHFR; Urlaub and Chasin, 1980), an essential enzyme which catalyses the conversion of folate to tetrahydrofolate (a precursor in DNA biosynthesis). CHO DHFR$^-$ cells have been extensively characterised as fast growing and can be easily transfected. Also CHO DHFR$^-$ cells allows the use of SV40-DHFR chimaeric plasmids (described above) to select for cells able to grow in the absence of thymidine, glycine and hypoxanthine. Moreover, these cells adapt well to growth on monolayer and in suspension and have been used successively in a variety of production systems such as human thyroid stimulating hormone (rh TSH; Lubiniecki *et al.*, 1989) and tissue plasminogen activator (tPA; Cole *et al.*, 1989). Finally, stable CHO cell lines containing a *dhfr* expression vector can be amplified to achieve increased gene copy number by culturing these cells in the presence of methotrexate (MTX), a selective agent that is toxic to cells. Cells overcome the toxic effects of this selective agent by producing more DHFR as a result amplifying the cellular gene encoding DHFR (Bebbington and Hentschel, 1987). Subsequently large chromosomal regions (hundreds to thousands of kb) are amplified (as described in section 1.2.2.3). The strategy involved in designing a CHO cell line for targeted integration of transgenes is described in the next section.

5.2 Design of Cre-loxP mediated targeted integration of transgenes in CHO cells

To investigate the Cre-loxP system as a tool for genomic targeting in CHO cells, a loxP site was initially placed into the chromosome of the CHO *dhfr*$^-$ cell line (CHO-DUKX B11) via a plasmid construct containing a reporter gene, Secreted Alkaline Phosphatase (SAP) and a *dhfr* selectable marker gene. The plasmid integrates randomly into CHO chromosome. The integration event can occur anywhere in the chromosome, either at a promoter region, an enhancer/repressor region, a non-coding region or a region in the chromosome under cell-cycle/tissue specific induction. Integration into any of these regions would have a direct effect of the level of gene expression, in this case the reporter gene, SAP. Thus a low SAP activity would indicate integration into a region in the chromosome which does not favour gene expression and a high SAP activity would indicate integration into a region of the chromosome favourable for gene expression. Because a loxP site is placed in the same plasmid containing the SAP gene, the level of SAP activity would be used as an indicator in screening a large number of cell lines each containing a single copy

insertion of the plasmid in order to isolate cell lines in which the loxP site has been placed in a region of the chromosome favourable for gene expression.

To develop a positive selection for the site-specific integration event two approaches were taken: In the first approach, a promoterless *lac*Z/neo^r fusion construct containing a loxP site was either co-transfected with a plasmid containing the *cre* gene or recombinant Cre protein purified from the baculovirus/insect system. Cre-mediated site-specific recombination targets the input DNA to a chromosomal loxP site which lies between the β-actin promoter and the SAP gene. Proper targeting should activate the expression of the bifunctional *lac*Z/neo^r fusion gene from the chromosomally placed β-actin promoter. In addition, transcription of the SAP reporter gene was under the control of the β-actin promoter, thus the Cre-mediated recombination event should not only give rise to cells resistant to the neomycin analogue G418 and expressing β-galactosidase activity but also loss or reduction of SAP activity (Figure 5.1).

In the second approach, a promoterless CAT gene construct containing a loxP site was co-transfected with a plasmid containing the *cre* gene. Similarly, the Cre-mediated integration of the promoterless CAT gene to a chromosomal β-actin promoter-loxP site should activate the expression of a functional CAT gene from the chromosomally placed β-actin promoter. In the case of targeting the promoterless CAT gene, the result of the targeting event should give rise to cells expressing CAT activity and the loss or reduction of SAP activity.

5.3 Results

The following sections describes the construction of CHO-DUKX B11 cell line incorporating the Cre-loxP site-specific recombination system as a tool that would allow the reproducible high level expression of recombinant proteins.

5.4 Construction of pSAP-RDN/loxP

As described in the previous section, Cre-mediated site-specific recombination in CHO cells requires a chromosomal loxP target for the integration event to occur. In order to develop a positive selection for the integration of loxP site(s) into CHO

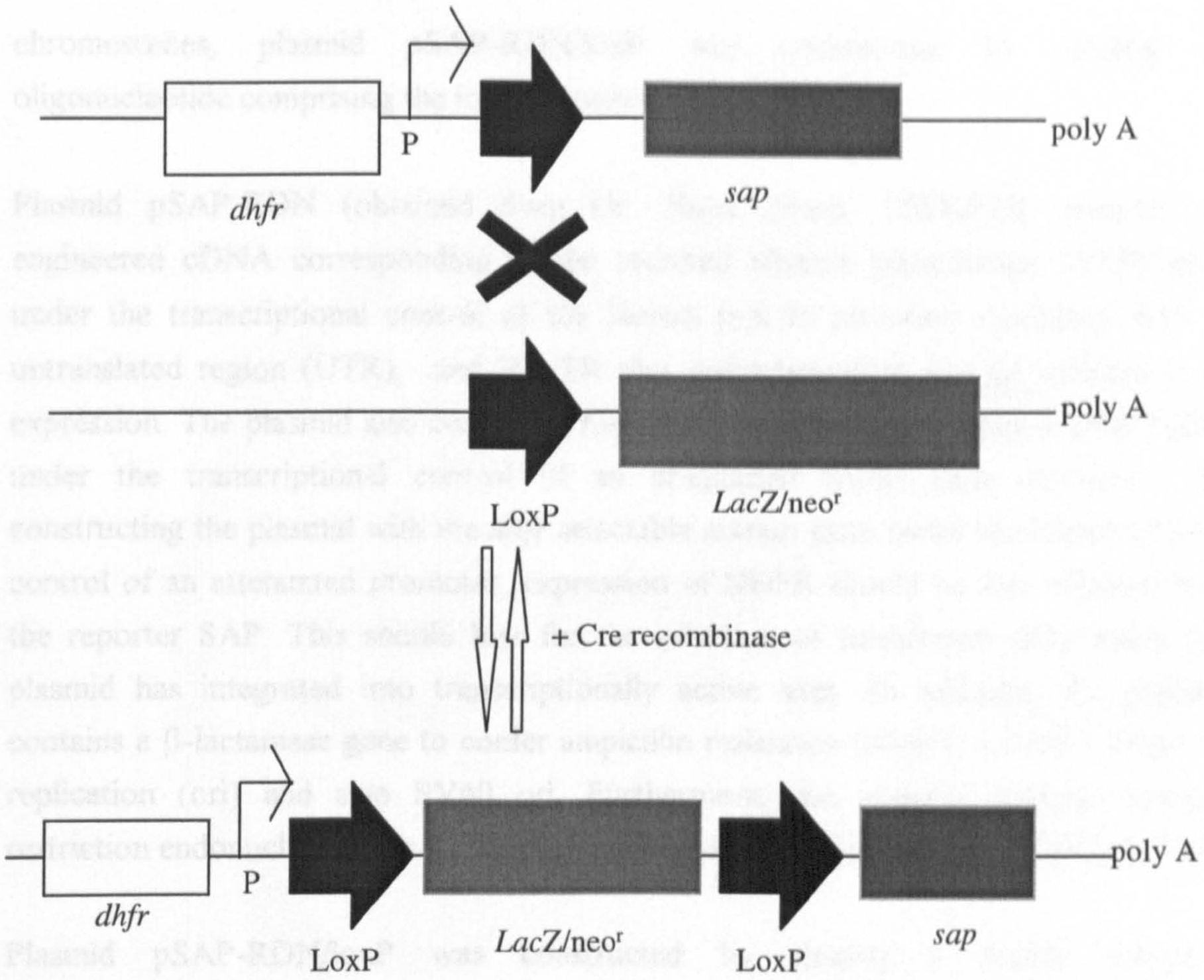


Figure 5.1-Schematic representation of Cre-mediated targeting of a an exogenous DNA to a chromosomal loxP site.Initially a loxP site was placed in the genome of CHO DHFR$^-$ cells via a plasmid encoding a dihydrofolate reductase gene (*dhfr*) and a secreted alkaline phosphatase (*sap*) reporter gene under the transcriptional control of β–actin promoter (P). Cre recombinase mediated targeting of a promoterless *lacZ*/neo^r fusion to the chromosomal loxP site activates the expression of the fusion gene from the β-actin promoter initially driving the SAP expression. This results in targeted clones which resistant to the neomycin analogue G418, positive for β-galactosidase expression and give loss or reduction in SAP expression.

chromosomes, plasmid pSAP-RDN/loxP was constructed by cloning an oligonucleotide comprising the loxP sequence into pSAP-RDN.

Plasmid pSAP-RDN (obtained from Dr. Nigel Sharp, GWR&D), encodes an engineered cDNA corresponding to the secreted alkaline phosphatase (SAP) gene under the transcriptional control of the human β-actin promoter combined with 5' untranslated region (UTR), and 3' UTR plus polyadenylation site for efficient gene expression. The plasmid also contains a functional dihydrofolate reductase (*dhfr*) gene under the transcriptional control of an attenuated SV40 early promoter. By constructing the plasmid with the *dhfr* selectable marker gene under the transcriptional control of an attenuated promoter, expression of DHFR should be less efficient than the reporter SAP. This should bias for the selection of transfected cells where the plasmid has integrated into transcriptionally active sites. In addition, the plasmid contains a β-lactamase gene to confer ampicillin resistance (Amp^r), a ColE1 origin of replication (ori) and also SV40 ori. Furthermore, the plasmid contains various restriction endonuclease sites for cloning and restriction digestion analysis (Figure5.2).

Plasmid pSAP-RDN/loxP was constructed by cloning a double stranded oligonucleotide comprising the loxP sequence into the unique *Hind*III site in pSAP-RDN. In order to bias for the ligation of the loxP sequence and the vector, pSAP-RDN was dephosphorylated (see section 2.2.9) after digestion with the *Hind*III restriction endonuclease. Results of the ligation event was analysed by transforming ligated DNA fragments into *E. coli* DH5α competent cells and transformed cells observed as ampicillin resistant colonies on L-agar plate (see section 2.2.13). The results showed fifteen colonies from the control plate (corresponding to self-ligated vector DNA) whilst forty-three colonies were observed as a result of pSAP-RDN and loxP ligation (test plate). This observation suggests that every 2 out of three colonies observed on the test plate should correspond to vector plus insert ligation.

Out of ten colonies (corresponding to pSAP-RDN + loxP ligation) which were analysed by sequencing only four had the loxP sequence inserted into pSAP-RDN. The sequence analysis further showed that the loxP site had inserted into pSAP-RDN in two orientations as determined by the directionality of the spacer region. In the one orientation loxR (Figure 3a), the loxP sequence was observed to have two ATG translational initiation codons upstream of the translational initiation codon of the SAP gene, with the first ATG codon out of frame with the SAP gene. Furthermore, a purine (A) base was observed at position -3 of the first ATG codon in the loxP

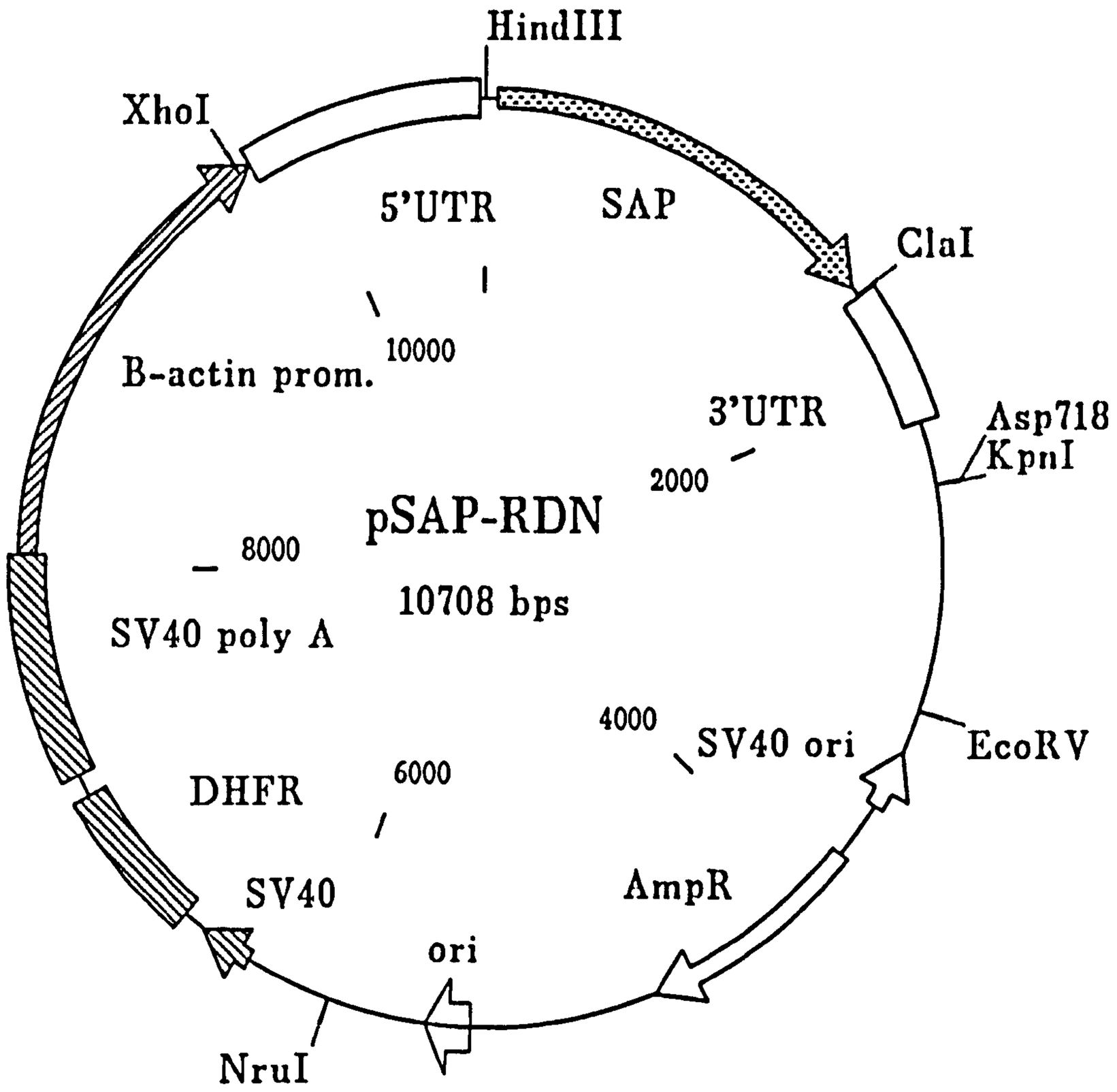


Figure 5.2-Map of pSAP-RDN.

Plasmid pSAP-RDN encodes a functional dihydrofolate reductase (*dhfr*) gene under the transcriptional control of attenuated SV40 early promoter. The plasmid also contains an engineered cDNA corresponding to the secreted alkaline phosphatase (SAP) gene under the transcriptional control of the human β-actin promoter combined with 5' untranslated region (UTR), and 3' UTR (plus polyadenylation site). In addition the plasmid contains a β-lactamase gene to confer ampicillin resistance (amp^r), a ColE1 origin of replication (ori) and SV40 ori.

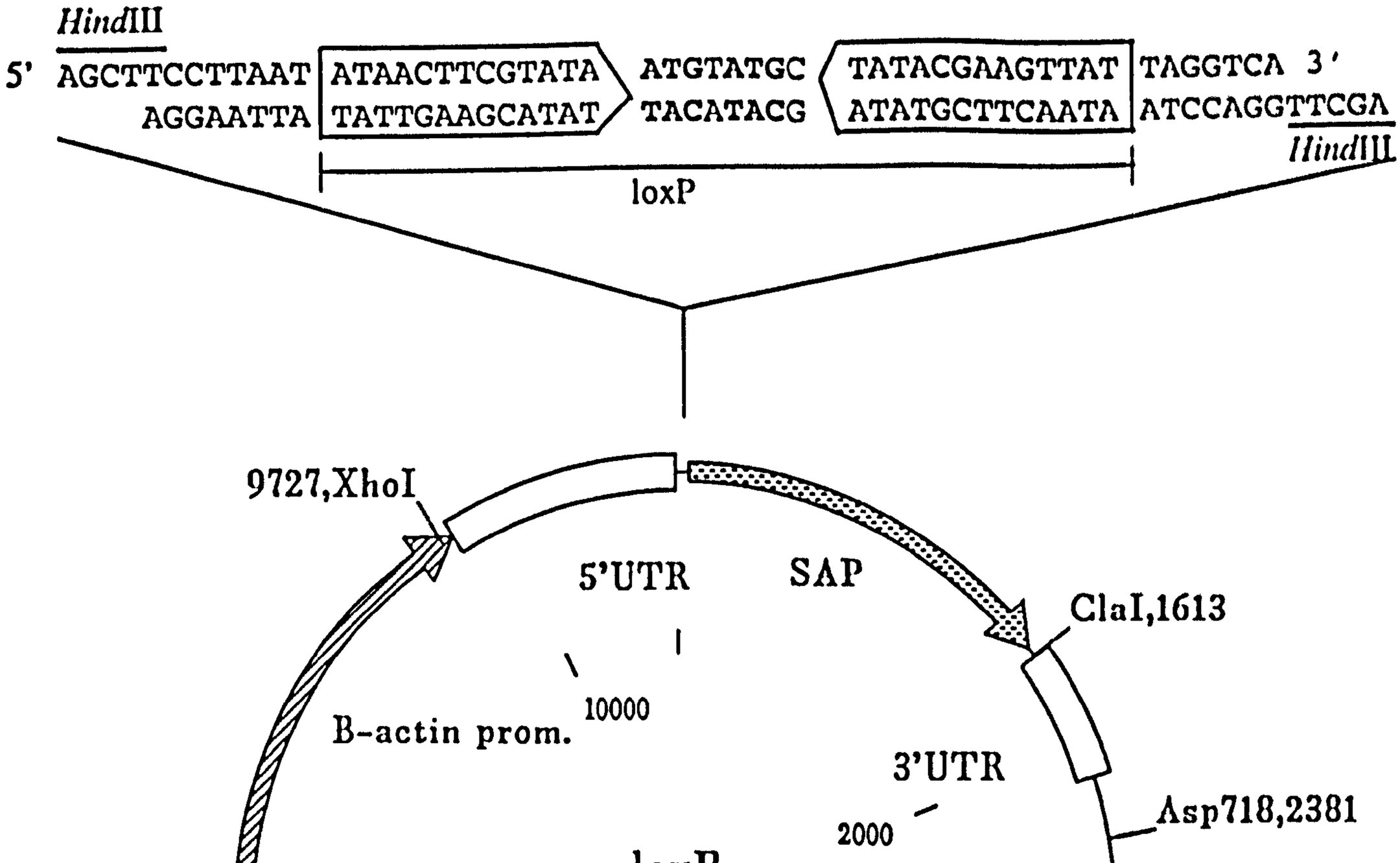


Figure 5.3a- Map of plasmid loxR

A double stranded oligonucleotide sequence corresponding to the loxP site (which consists of two 13 bp inverted repeats [shown in box with arrowheads] flanking an 8 bp spacer region) with *Hind*III recognition ends was cloned into the unique *Hind*III site in pSAP-RDN (see Figure 5.2) to generate loxR. The loxP site in this orientation was observed by sequencing analysis to have two ATG translational initiation codons (shown in bold) upstream with respect to the first ATG codon of the SAP gene.

sequence suggesting that the ATG codon is in an appropriate context for translational initiation (Kozak, 1987). Translational initiation from the ATG in the loxP sequence instead of the authentic ATG in the SAP gene could inhibit the transcription of the SAP reporter gene and therefore decrease SAP expression. However no such ATG translational initiation codons were observed when the loxP sequence was inserted in the other orientation, loxL (Figure 3b).

To construct a recombinant CHO cell line containing loxP site(s); pSAP-RDN/loxP (with the loxP site in both orientations) were independently transfected into CHO-DUKX B11 cell line as described in the next section.

5.5 Construction of CHO cell lines containing a single loxP site by lipofection

A loxP site was introduced into CHO chromosome by transfecting CHO-DUKX B11 cells with the plasmid pSAP-RDN/loxP (with the loxP sites in both orientation; i.e., loxL and loxR). The plasmid contains a *dhfr* gene to allow for the selection of transfected cells that have stably integrated the loxP plasmid DNA by growing transfected cells in medium lacking hypoxanthine and thymine.

Circular plasmid DNA was transfected into CHO-DUKX B11 using lipofection. In this method, a lipofectam reagent which contains a spermine group, with a strong affinity for DNA, is covalently attached to a lipid moiety. The trivalent positively charged spermine headgroups bind to the negatively charged DNA, and in the presence of excess lipopolyamine, cationic lipid-coated DNA particles are formed. The lipid allows association of the particles with the cell membrane and internalisation of the DNA probably occurs by endocytosis.

CHO-DUKX B11 (1×10^6) cells were transfected separately with 16 µg of loxL and 16 µg of loxR mixed with 80 µg of Transfectam. Two days post transfection, cells were plated at six dilutions; 100, 50, 10, 5, 1, 0.33 cells/well (described in 2.4.3) into DHFR select medium. Approximately 400 $DHFR^+$ colonies/10^7 transfected cells were obtained twenty days post transfection. It should be noted that this transfection efficiency is an underestimate because wells that contained more than one colony were disregarded. Furthermore, the calculation is based on the assumption that no new colonies appeared after 20 days. Plasmids loxL and loxR both contains the reporter (SAP) gene, therefore $DHFR^+$ clones isolated after transfection can also be assayed

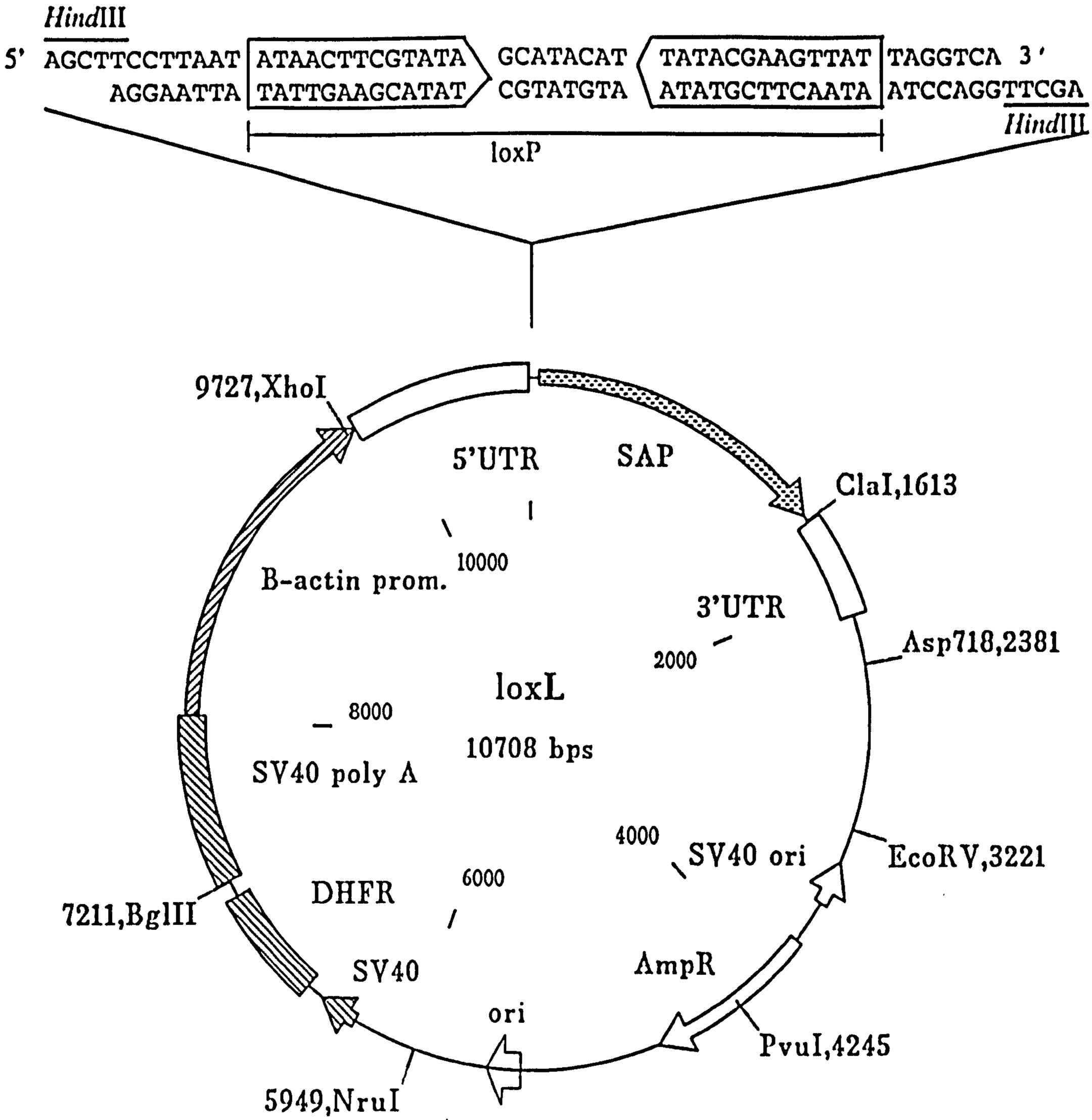


Figure 5.3b- Map of plasmid loxL

A Double stranded oligonucleotide sequence corresponding to the loxP site (which consists of two 13 bp inverted repeats [shown in box with arrowheads] flanking an 8 bp spacer region) with *Hind*III recognition ends was cloned into the unique *Hind*III site in pSAP-RDN (see Figure 5.2) to generate loxL. The loxP site in this orientation does not contain any ATG codon(s).

for SAP activity in order to assess the transcriptional level of SAP gene in independent cell lines and to select for cell lines in which the loxP site has been placed in a region of the chromosome favourable for gene expression. The analysis of SAP expression in DHFR$^+$ clones is described in the next section.

5.5.1 Assay for SAP activity in DHFR$^+$ clones

Many genes have been used for monitoring gene expression in eukaryotic cells. This includes β-galactosidase (An *et al.*, 1982), xanthine-guanine phosphoribosyl transferase (Chu *et al.*, 1985) and thymidine kinase (Searle *et al.*, 1985). However, all of these genes have disadvantages, such as the presence of a similar endogenous enzyme activity or the requirement for a complicated assay. Another reporter gene that has become widely used is the *E. coli* enzyme chloramphenicol acetyl transferase (CAT; Gorman *et al.*, 1982). The major advantages of this enzyme is the lack of any equivalent enzymatic activity in eukaryotic cells and the existence of highly sensitive, radioactivity-based assays for CAT. The assay is however expensive to do and may require the use of radioactivity. SAP, the secreted form of human placental alkaline phosphatase on the other hand avoids several of these disadvantages, for instance, it is far simpler and less expensive to use.

In addition, SAP has several advantages over the membrane bound form of alkaline phosphatase (which is frequently assayed in diagnostic laboratories); (i) SAP is not significantly affected by heating at 65^{o}C for 5 minutes (Stigband, 1984; Berger *et al.*, 1987). The resistance of SAP to heat treatment, which effectively eliminates other AP activities is indicative of the high stability of SAP. (ii) Because SAP is expressed in the culture medium, other kinetic experiments can be done using the same culture medium, through sequential sampling of the medium. In addition, the lack of an extraction step removes the possibility of trapping AP activity in the discarded particulate material and also segregates the indicator SAP gene away from any endogenous cell surface AP activity. Consequently, the cells can be used for other purposes, such as RNA preparation.

The SAP assay used (detailed in section 2.8.3) is essentially the standard colorimetric assay for AP (McComb and Bowers, 1972) modified by the addition of a 65^{o}C pre-heating step to the reaction mixture. The increase in absorbance at 405nm, which results from the hydrolysis of ρ-nitrophenolphosphate to ρ-nitrophenol, is proportional to the AP activity (McComb and Bowers, 1972). SAP activity (N. Sharp, GWR&D,

pers. comm.) can be accurately quantified in a small volume permitting the use of any of several available automoatic 96-well plate ELISA readers for the assay.

Forty DHFR$^+$ clones isolated from each transfection were assayed for SAP activity. The level of SAP expression from loxR transfected cells were observed to be between 10-100 fold lower compared to the level of SAP expression from loxL transfected cells. Because DHFR$^+$ clones isolated from loxR transfection contains two ATG codons (in the loxP sequence) upstream from the authentic ATG codon of the SAP gene, this result suggests that in stably transfected cells containing loxR, the ribosome may recognise and initiate translation from the ATG codon in the loxP sequence instead of the authentic ATG codon of SAP. Consequently translation of the SAP gene is repressed.

Most of the DHFR$^+$ clones isolated from loxR transfection showed similar and very low levels of SAP activity which, as mentioned previously, could be due to translational initiation from the first ATG codon in loxR which results in a decrease SAP expression. However, a good representation of clones expressing different levels of SAP were required in order to identify those clones which had integrated the loxP site(s) in chromosomal positions favourable for gene expression. Therefore most of the these clones were disregarded for further analysis. On the other hand, DHFR$^+$ clones isolated from loxL transfection showed high and relatively different levels of SAP expression, hence were propagated for further analysis. In addition, four of the DHFR$^+$ clones isolated from loxR transfections were also propagated to investigate if the different SAP expression level observed from the two different plasmid transfections were due other factor(s) such as differing copy number of the plasmid integrated and/or different chromosomal position effects in the two separate transfection events.

The initial SAP activity was used to identify cell lines expressing SAP. In order to characterise these cell lines and accurately get a representation of the level of SAP expressed in these cell lines, the specific production rate (S.P.R; section 2.8.4) of SAP was determined for each clone. Results shown in Table 5.1 confirmed the initial observation that the level of SAP expression from transfected cells containing loxR were much lower compared to transfected cells containing loxL. It is possible that the different expression levels of SAP observed between the different transfected cells is due to differing copy numbers of the loxL and loxR plasmid integrated and also different positions of integration into the CHO genome. In order to investigate the

pSAP-RDN/loxP (loxL)	SAP (arbitrary units/10^6 cells/day)	pSAP-RDN/loxP (loxR)	SAP (arbitrary units/10^6 cells/day)
2-12	1.1	1-1	0.9
2-20	2.3	1-3	0.05
2-11	2.5	1-15	0.03
2-29	4.7	1-20	0.01
2-26	15.2		
2-9	14.8		
2-19	20.8		
2-16	22.1		
2-2	29.7		
2-21	31		
2-24	40		
2-1	81		
2-18	100		
2-6	100		
2-8	115		

Table 5.1- Comparison of standard production rate (SPR) of SAP expression from independent CHO-DUKX B11 cells transfected with recombinant pSAR-RDN/loxP (where the loxP site is cloned in one of two orientations [L or R]).

One million cells from transfected CHO-DUKX B11 cells was grown in 5 ml of DHFR$^+$ select medium in tissue culture flask overnight. The medium was replaced with fresh medium. 24 hours after the addition of the fresh medium (± 30 min), 0.5 ml of the medium was removed to determine SAP activity (described in section 2.4.7). The adherent cells in the flask were detached by versene and trypsin treatment and viable cell density determined. Arbitrary unit is defined as the change in absorbance at A_{405nm} in 30 minutes (readings were done in duplicates).

$$\text{S. P. R. (arbitrary units/}10^6\text{ cells/24 hours)} = \frac{\text{Arbitrary units of SAP activity}}{\text{cell density (}10^6\text{/ml)}}$$

copy number of loxL and loxR integrated in each transfection, a Southern hybridisation was done, described in the next section.

5.5.2 Analysis of loxP copy number by Southern hybridisation

For Cre-mediated site-specific targeting of exogenous DNA into CHO-DUKX B11 cells, a single loxP sequence placed in a unique site in the chromosome is required. This is because multiple loxP copies in the chromosome could lead to chromosomal translocation and rearrangements. In order to select for CHO transfected cell lines containing a single copy of the loxP plasmid and to assess if the copy number of plasmid integrated in CHO chromosome would correlate with different expression level of SAP, $DHFR^+$ clones were analysed by Southern hybridisation.

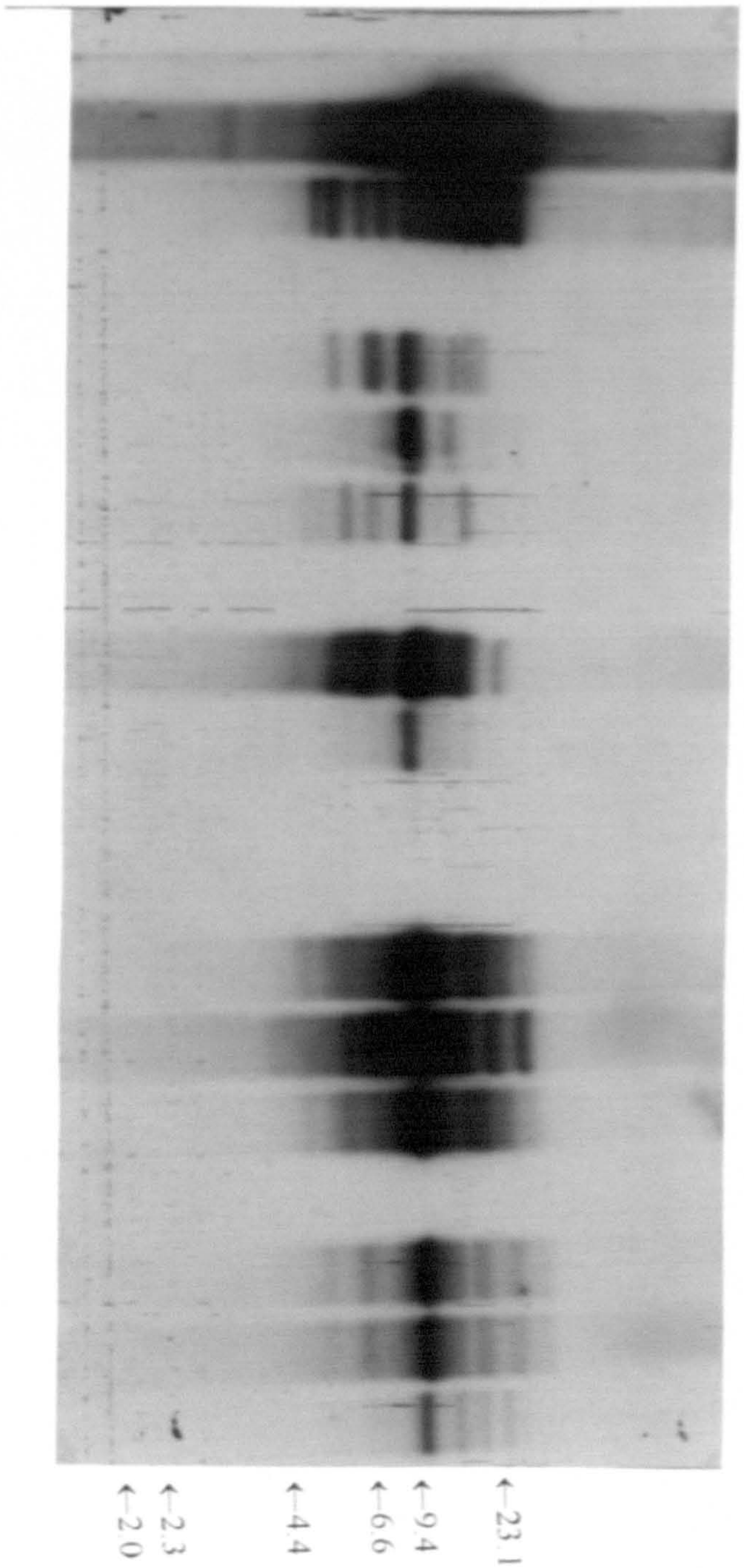
Genomic DNA prepared from CHO transfected cell lines (10 µg) was digested with *Eco*RV and separated by electrophoresis on agarose gel. The DNA was transferred onto nylon membrane and hybridised with an 800 bp radioactive probe corresponding to SV40 polyA fragment. The SV40 polyA fragment was obtained from pSV-SB (obtained from Nigel Sharp, GWR&D) by digestion with *Bgl*II and *Bam*HI. The CHO genome does not contain SV40 polyA whereas the transfected DNA (i.e. loxL and loxR) does. Therefore transfected cells which have integrated the plasmid should hybridise to this probe in Southern hybridisation. Southern hybridisation was visualised by autoradiography (Figure 5. 4).

The restriction enzyme *Eco*RV will cut the genomic DNA in many places, but the plasmid insert is only cut in one place. When the SV40 polyA signal is used as a probe in the Southern hybridisation, a genomic DNA fragment containing a single copy of loxP plasmid is expected to hybridise to the radioactive probe and appear on the autoradiograph as one signal. If however, the plasmid integrated as monomers in tandem repeats, then a 10.7 kb signal is expected. Multiple copy insertions into multiple chromosomal sites is expected to give rise to signals of variable sizes.

Results of the Southern blot analysis showed the presence of a very intense signal of about 10.7 kb co-migrating with the same electrophoretic mobility in all the clones analysed. In addition, other signals of varying sizes were observed. The 10.7 kb signal probably represents the full length of the loxP plasmid, suggesting that plasmid had integrated into CHO chromosome in tandem repeats, whilst the other signals indicates multiple copy integration into different chromosomal sites. The difference in the

Figure 5.4- Detection of pSAP-RDN/loxP transfection into CHO-DUKX B11 cells by Southern blot analysis.

Genomic DNA (10 μg) prepared from independent CHO-DUKX B11 transformants was digested with *Hind*III. Digested DNA was separated by electrophoresis on an agarose gel and DNA fragments transferred onto a nylon membrane and hybridised using a radioactive SV40 polyA fragment. Lanes are labelled according to clone number. DNA marker sizes to the right are in kilobases.


CHO B11
CHO B11 (amplified)
2-12
2-20
2-11
2-29
2-26
2-9
2-19
2-16
2-2
2-21
2-24
2-1
2-18
2-6
2-8
1-1
1-20
←2.0
←2.3
←4.4
←6.6
←9.4
←23.1

intensity of hybridisation between the different clones are the result of inaccuracy in the determination of DNA concentration which was initially observed by ethidium bromide staining of *Eco*RV digested genomic DNA in agarose gel (data not shown). No signals were observed in the CHO DUKX B11 (CHO B11) negative control lane which does not contain sequences homologous to the SV40 polyA fragment. Whereas the positive control, CHO B11 amplified cell line (transfected with an SV40 polyA containing plasmid, similar to the one used in this study) showed a smeared intense signal showing that the observed signals in the transfected clones corresponds to the integration of pSAP-RDN/loxP plasmid into CHO chromosome(s).

It is very difficult to correlate the copy number of the loxP plasmid with the level of SAP expression determined in each clone, because the different signals observed and also the varying intensities of the signals is hard to interpret. The result shows that lipofection method of transfection leads to multiple copy of the loxP plasmid inserted into multiple sites. Transfected cells containing multiple copies of the loxP sites inserted in multiple sites in the chromosomes cannot be used to investigate the subsequent site-specific integration event because this could lead to chromosomal translocation and rearrangements. For this reason, a different transfection method, electroporation (described in the next section) was used to investigate if this method might yield single copy transfected cell lines.

Because the ATG codons in the loxP sequence in loxR appeared to interfere with the expression level of SAP in CHO-DUKX B11 cells transfected with loxR, all subsequent investigations were done using plasmid loxL.

5.6 Transfection by electroporation

Electroporation is the process whereby DNA mixed with cells in suspension is exposed to a high-voltage electric field. This creates pores in the plasma membranes (Kinosita and Tsong, 1977) of treated cells that are large enough to allow the passage of macromolecules such as DNA into the cells. Such DNA molecules are ultimately transported to the nucleus, and a subset of these molecules are stably integrated into the host chromosome. The reclosing of the membrane pores is both time and temperature-dependent and thus is delayed by incubation at 0^{o}C, thereby increasing the probability that the molecule of interest will enter the cell. Linearised DNA appears more recombinogenic than supercoiled or nicked circular DNA (Chu *et al.*, 1987) and

thus leads to a higher percentage of stably transformed cells. Supercoiled DNA serves as a superior transcriptional template in transient transfection experiments. Transfection frequency also depends on a number of other factors; the most obvious possibilities include DNA concentration, electric field strength and pulse length, size of DNA or DNA complex, cell type, and condition of the cell membrane. Voltage is a critical parameter for electroporation and, for a given capacitance and buffer, there is a sharply defined voltage for optimal transfection efficiency (Chu *et al.*, 1987). At higher voltages the cells are irreversibly damaged and at lower field strengths the cells do not take up appreciable amounts of DNA.

Electroporation has a number of advantages over other methods of transfection such as calcium phosphate co-precipitation (Graham and van der Eb, 1973) and DEAE-dextran (McCutchan and Pagano, 1968) mediated DNA transfection. These advantages include: the successful electroporation of cells which are totally resistant to calcium phosphate or DEAE-dextran mediated gene transfer especially those cell lines which grow in suspension e.g. mouse erythroleukemia cells and other cells of hematopoietic origin (Boggs *et al.*, 1986); ease of operation; reproducibility of conditions; applicability to cells which are either adherent or those growing in suspension; utility for both transient and stable transfection procedures; and the capacity to control copy number of transfected DNA molecules.

5.6.1 Construction of CHO cell lines containing a single loxP plasmid by electroporation.

As described in the previous section electroporation provides another method for transfecting DNA into mammalian cells. Electroporation though efficient for transfection can also kill cells. Therefore the electroporation conditions had to be optimised for CHO-DUKX B11 cell line.

Two different sets of electroporation parameters that have been used for transfection in CHO cell lines (Godwin *et al.*, 1994; Fukushige and Sauer, 1992) were investigated for efficient transfection into CHO-DUKX B11 cell line. CHO-DUKX B11 (10^7) cells were transfected with 10 µg of pSAP-RDN/loxP by electroporation at either 960 µF and 250V or 500 µF and 450V. Five minutes after electroporation, cell viability was assessed by staining an aliquot of transfected cells with trypan blue. Cell death was not detected by the trypan blue staining immediately after electroporation. Instead, the cells adhered to the culture dish initially, only to die and float off into the medium over

the next 48 hours. Cell death was estimated to be 90% when electroporation parameters were 500 μF and 450V, whilst cell death was estimated to be about 60% when electroporation parameters were 960 μF and 250V. Cell viability of about 50% is required for efficient transfection, therefore all subsequent electroporations were done at 960 μF and 250V, with the assumption that 60% viability would be acceptable.

In the previous transfection using lipofectam, circular plasmid DNA was used. However, in this transfection by electroporation, linearised plasmid DNA was used. This is because previous experiments comparing the transfection (electroporation) efficiency of linearised to supercoiled plasmid DNA have demonstrated a 2-20 fold greater efficiency of stable transfection with linear plasmid (Chu *et al.*, 1987).

CHO-DUKX B11 (10^7) cells were transfected with either 10 μg or 1 μg of loxL linearised at the *Eco*RV site. Electroporated cells were plated out in complete medium onto T-75cm^2 flasks. Forty eight hours post transfection, electroporated cells were resuspended into select medium and dilution cloned to isolate individual clones. Transfected cells were plated at six dilutions; 100, 50, 10, 5, 1, 0.33 cells/well (as described in section 2.4.3) into DHFR selection medium.

Transfection of 10 μg of linearised pSAP-RDN/loxP plasmid produced 158 $DHFR^+$ colonies/10^7 transfected cells. A somewhat lower number of $DHFR^+$ colonies, 120/10^7 transfected cells were obtained by transfecting 1 μg of the linearised plasmid. $DHFR^+$ clones were propagated for further analysis.

During the propagation step to isolate stable cell lines, most of the colonies did not survive selection resulting in fewer colonies that are $DHFR^+$; 40 clones survived from the transfection with 10 μg DNA, whilst only 13 clones survived from the transfection using 1 μg of DNA. The colonies that disappeared, probably survived initially by expression of non-integrated episomal (self ligated vector DNA) plasmid DNA and then died after the DNA was degraded or diluted out by cell division, because the plasmid does not replicate autonomously. Such "false" positive clones have been reported by others (Kathryn Bentley, WRL, pers. comm.). In addition, many colonies picked at late times (from 20-30 days post-transfection) grew very slowly and were discarded.

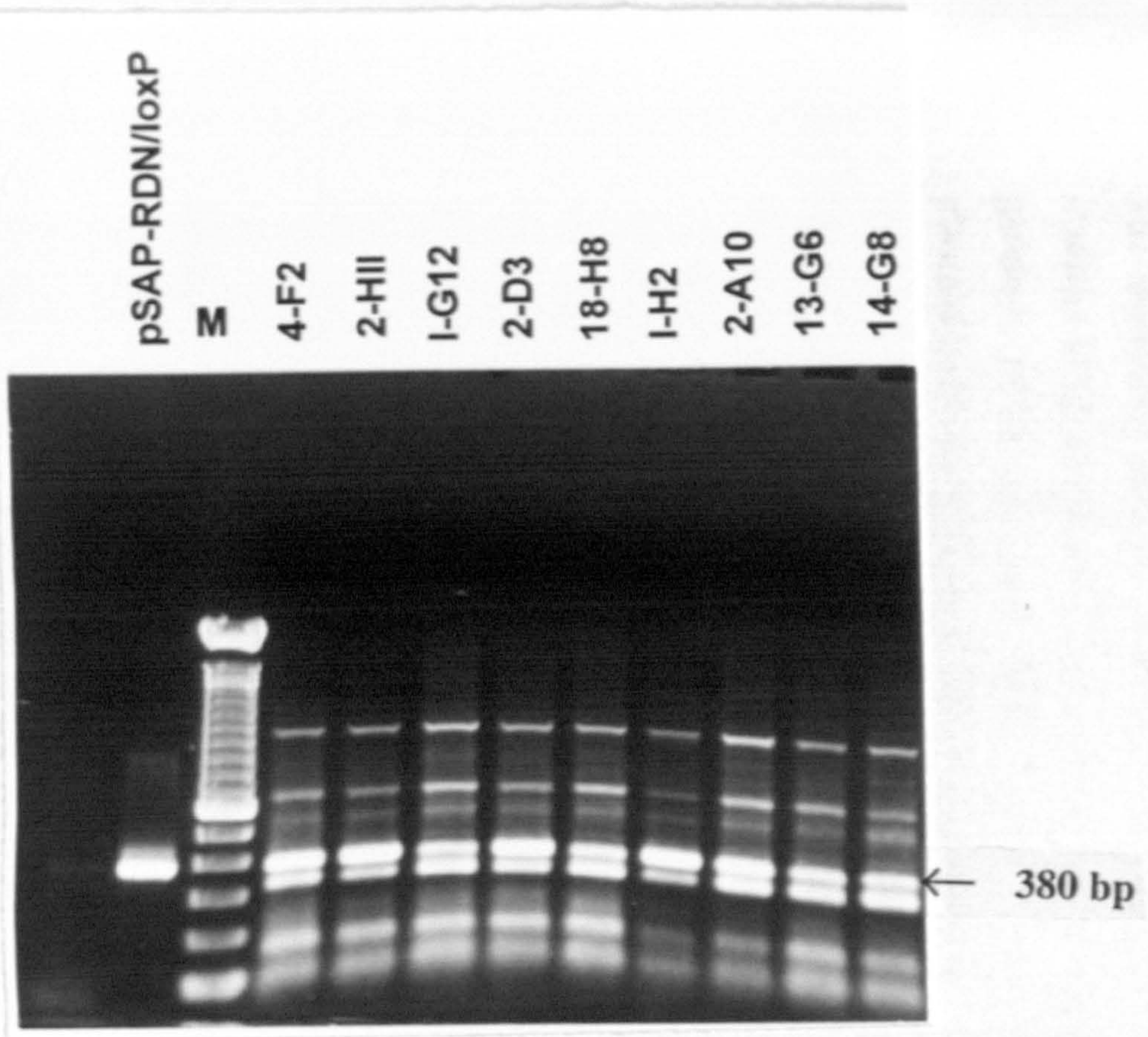
A quick SAP assay was done to identify those $DHFR^+$ clones expressing SAP. Results showed different level of SAP expression in all of the $DHFR^+$ clones analysed (data not shown). However, the difference in the level of SAP expression could be due to

differing copy number of the loxL plasmid integrated into CHO chromosome and also from chromosomal position effects. Therefore further characterisation of the DHFR$^+$ clones was necessary to identify cell lines containing single copy insertion of the loxL plasmid.

5.6.1.1 Investigating DHFR$^+$ clones by polymerase chain reaction (PCR)

A functional chromosomal loxP site is required for the site-specific integration event to occur. Therefore a PCR analysis was done to determine if the DHFR$^+$ clones had integrated an intact loxP site. A combination of oligonucleotide primer sequences, designed as forward primer (CB1736- GCTATTCTCGCAGGATCAGTCGACC) to anneal within β-actin 5' UTR region and the reverse primer (CB1738- CCAGTTTGTCCTTCTTCTGCCCTTT) to anneal within the *SAP* gene in loxL (see Figure 5.3), were used to amplify sequences from genomic DNA prepared from DHFR$^+$ clones. If the loxP site is intact in the CHO DHFR$^+$ clones, a PCR analysis using these primers is expected to yield a 380 bp product.

Results (presented in Figure 5.5) shows a distinct band of 380 bp (indicated by an arrow) in all the DHFR$^+$ clones analysed. This band comigrated with the same electrophoretic mobility as the amplified DNA from pSAP-RDN/loxP plasmid (positive control) suggesting that all the DHFR$^+$ clones analysed contained the fragment (i.e. β-actin 5' UTR-loxP-SAP) of loxL plasimd intact. In addition, a sharp band which is slightly smaller than the 380 bp fragment were observed in all the clones analysed. This band was also present in the negative controls (CHO DNA and CHO amplified DNA, which do not contain the loxL plasmid) but not in the positive control (pSAP-RDN/loxP) showing that the primer sequences also amplified other genomic sequences. It is difficult to explain the strong signal of this band, however the absence of the 380 bp band in the two negative controls clearly shows that the 380 bp band is the result of the loxL plasmid integrated into the CHO genome. Other non-specific bands were also observed again suggesting that the primer sequences are also amplifying other genomic DNA sequences. The result of the PCR analysis although it gives an indication of the structure of the loxL plasmid in CHO genome, does not however, give any indication of the number of loxL plasmid(s) integrated. Therefore a Southern hybridisation was done to investigate the copy number which is described in the next section.



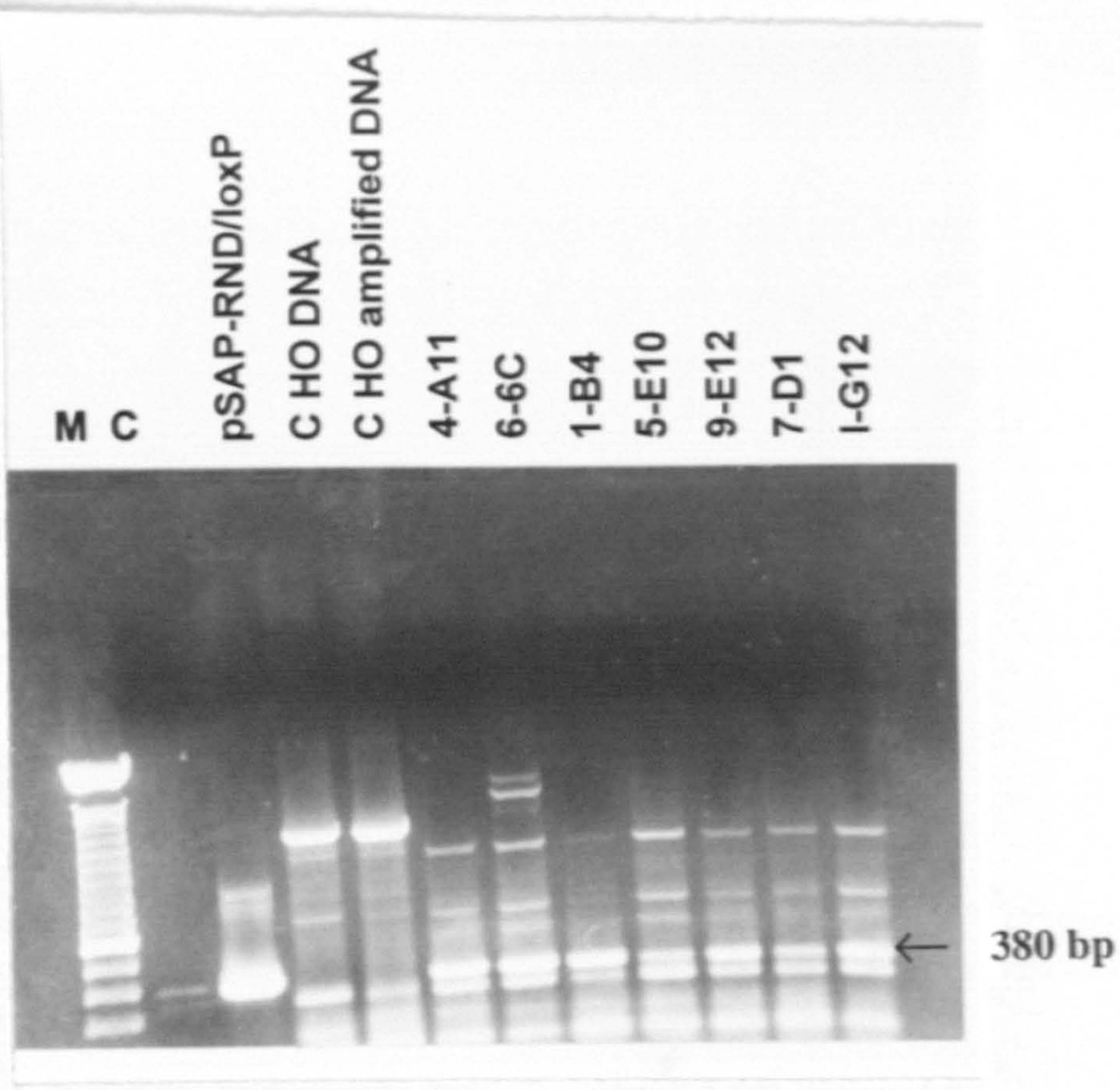


Figure 5.5- Agarose gel separation of PCR reactions of independent pSAP-RDN/loxP transfected CHO B11 cells.

PCR reactions were done using genomic DNA prepared from independent CHO-DUKX B11 transformants and a pair of primers corresponding to sequences within the β-actin intron region and the SAP gene. A 380 bp (indicated by an arrow) fragment was amplified from all the transfected cells analysed, the same size fragment as the positive control (pSAP-RDN/loxP. The 380 bp fragment was not observed in the negative controls; CHO-DUKX B11 or the CHO-DUKX B11 (amplified) cell line. Lanes are labelled according to the clone number. M indicates 100 bp DNA marker and C indicates negative control (primer only in the reaction).

5.6.1.2 Analysis of loxP copy number by Southern hybridisation

As previously mentioned, for Cre-mediated site-specific integration, it is preferable for the host cell line to contain a single copy of loxP site. In order to isolate DHFR$^+$ clones containing single copy insertions of the loxP site and also gain some insight into structural features of the integrated DNA, genomic DNA prepared from DHFR$^+$ clones was analysed by Southern hybridisation.

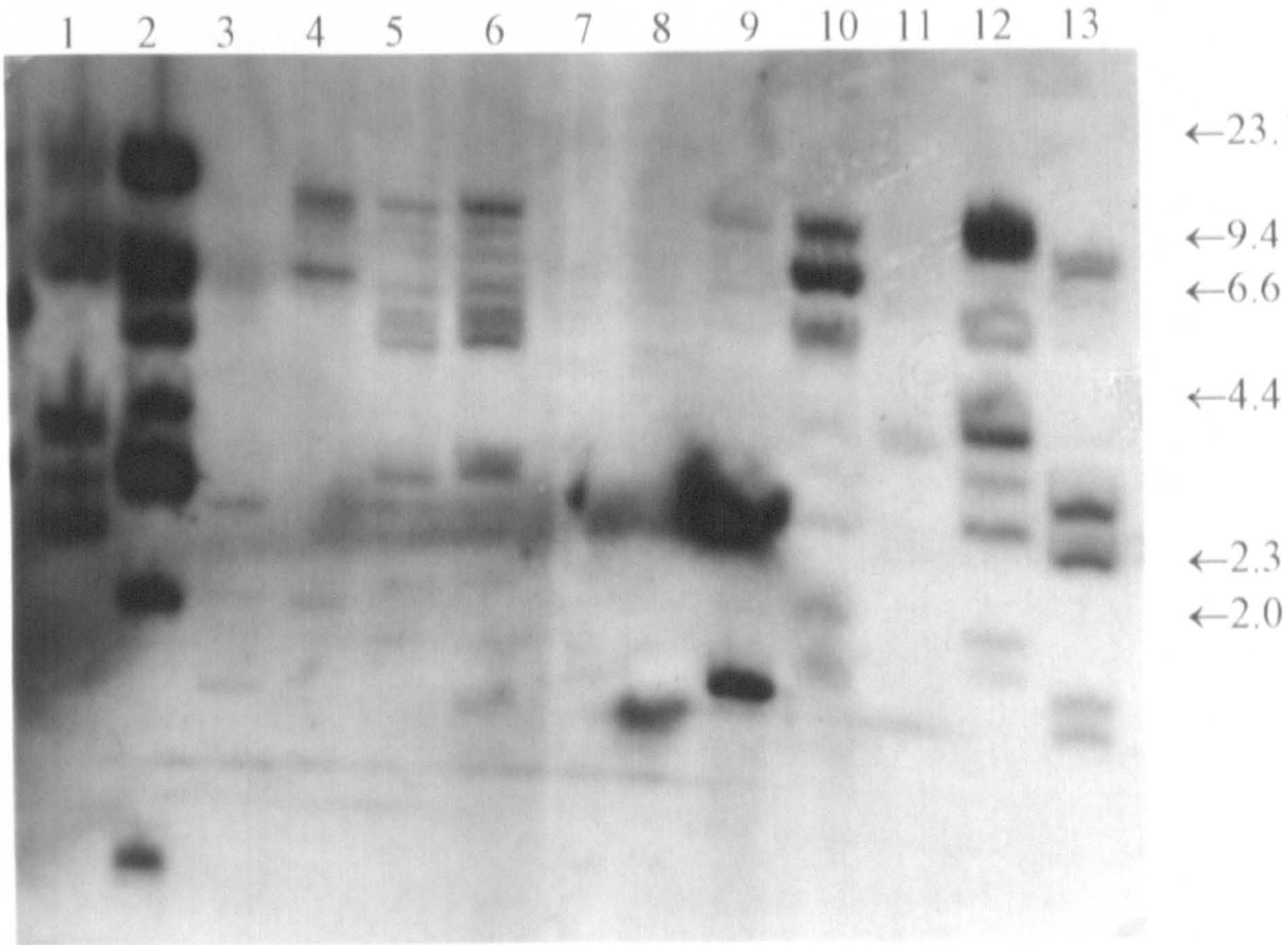
Genomic DNA (10 μg) prepared from individual DHFR$^+$ clones was digested with *Bgl*II (instead of *Eco*RV; see section 5.4.2) and separated by electrophoresis on agarose gels. [Genomic DNA was digested with *Bgl*II because the loxL plasmid was linearised by digesting the plasmid with *Eco*RV before transfection, therefore digesting the genomic DNA with *Eco*RV will probably just recover a linear loxL fragment which will not give much information as to whether integration occurs as multiple copies to multiple chromosomal regions]. The DNA was subsequently transferred onto nylon membranes and hybridised with a radioactive SV40 polyA probe (same probe used in section 5.4.2).

The restriction enzyme *Bgl*II will cut the genomic DNA in many places, but the inserted DNA is cut only once. When the SV40 polyA signal is used as a probe with genomic DNA from the DHFR$^+$ clones, the DNA fragment containing the loxP plasmid hybridises to the SV40 polyA probe and appears on the autoradiograph as a signal of 6.7 Kb plus an unpredictable amount of genomic DNA if the plasmid was integrated as a single copy. If two tandem copies of the plasmid are inserted at a single site in a head-to-tail orientation, then the restriction endonuclease site used to linearise the input DNA would remain intact. This event would be due to ligation of vector DNA via cohesive ends prior to integration. In this case restriction digestion with *Bgl*II will yield a unit length of 10.7 Kb and 4.1 Kb plus an unpredictable amount of genomic DNA. If two copies are integrated head-to-head or tail-to tail then *Bgl*II digestion would yield a 13.4 Kb fragment. However larger numbers of separate and or random integrations would give rise to multiple bands with varying intensities with or without degradation or rearrangement prior to integration.

Results of the Southern hybridisation from 10 μg pSAP-RDN/loxP transfection showed multiple signals in all the clones analysed, suggesting multiple copy number of the plasmid integrated into different chromosomal regions (Figure 5.6). When 1 μg of linearised pSAP-RDN/loxP was used in the Southern hybridisation analysis the

Figure 5.6- Detection of pSAP-RDN/loxP (10 μg) transfection into CHO B11 cells using electroporation.

Genomic DNA (10 μg) prepared from CHO-DUKX B11 cells transfected with pSAP-RDN/loxP was digested with *Bgl*II. The digested DNA was separated by electrophoresis on an agarose gel, and DNA fragments transferred onto a nylon membrane and probed with a radiolabelled SV40 poly A fragment. Lanes 1 to 13 corresponds to independent clones. DNA marker sizes to the right are in kilobases.


1 2 3 4 5 6 7 8 9 10 11 12 13
←23.1
←9.4
←6.6
←4.4
←2.3
←2.0

results showed that out of six clones analysed, two clones (namely 5-E10 and 1-G12) appeared to be single copy integrants and two other clones (namely 18-H8 and 14-G8) appeared to have the loxP plasmid in tandem repeat (Figure 5.7a). In the negative control experiment where genomic DNA from CHO-DUKX B11 cells used (which did not contain the SV40 polyA sequence) no signals were observed. However in the positive control experiment where CHO amplified DNA (which has been amplified to contain many copies of SV40 polyA sequence) a very strong smeared signal was observed. To verify further that these clones contained a single and tandemly repeated loxP sites respectively, genomic DNA prepared from the four clones were digested with *Xho*I and analysed alongside *Bgl*II digested DNA by Southern hybridisation (Figure 5.7b). The result provided further evidence that each of the two clones, 5-E10 and 1-G12 indeed contained a single copy insertion of the loxP plasmid, whilst each of the other two clones, 18-H8 and 14-G8 contained the loxP plasmid in tandem repeats.

The results from the Southern hybridisation analysis shows that only two out of thirteen clones had integrated a single copy of the loxL plasmid. However, the aim of this work was to isolate a large number of cell lines each containing a single copy insertion of the loxP site in order to identify cell lines which had integrated a lox plasmid in a chromosomal region which will be favourable for gene expression. Therefore another transfection was done in the hope of isolating more cell lines containing a single copy insertion of the loxL plasmid.

5.6.2 Further transfection by electroporation to isolate CHO cell lines containing a single loxP plasmid

CHO-DUKX B11 (1×10^7) cells were transfected with 1 μg of *Eco*RV linearised loxL plasmid by electroporation (as described above). Transfected cells were selected, propagated to establish stable cell lines and assayed for SAP expression as described previously. Eleven $DHFR^+$ clones which were positive for SAP expression were subsequently analysed by Southern hybridisation. Southern hybridisation was done as described in the previous section (Figure 5.8).

Results showed that one clone (namely 13-G6) out of the eleven clones analysed had a single copy of the loxP plasmid integrated into the chromosome. Whilst the other clones showed multiple signals to suggest multiple copy integrations into multiple chromosomal sites. However, clone 7-D1 which showed one strong signal and a weak

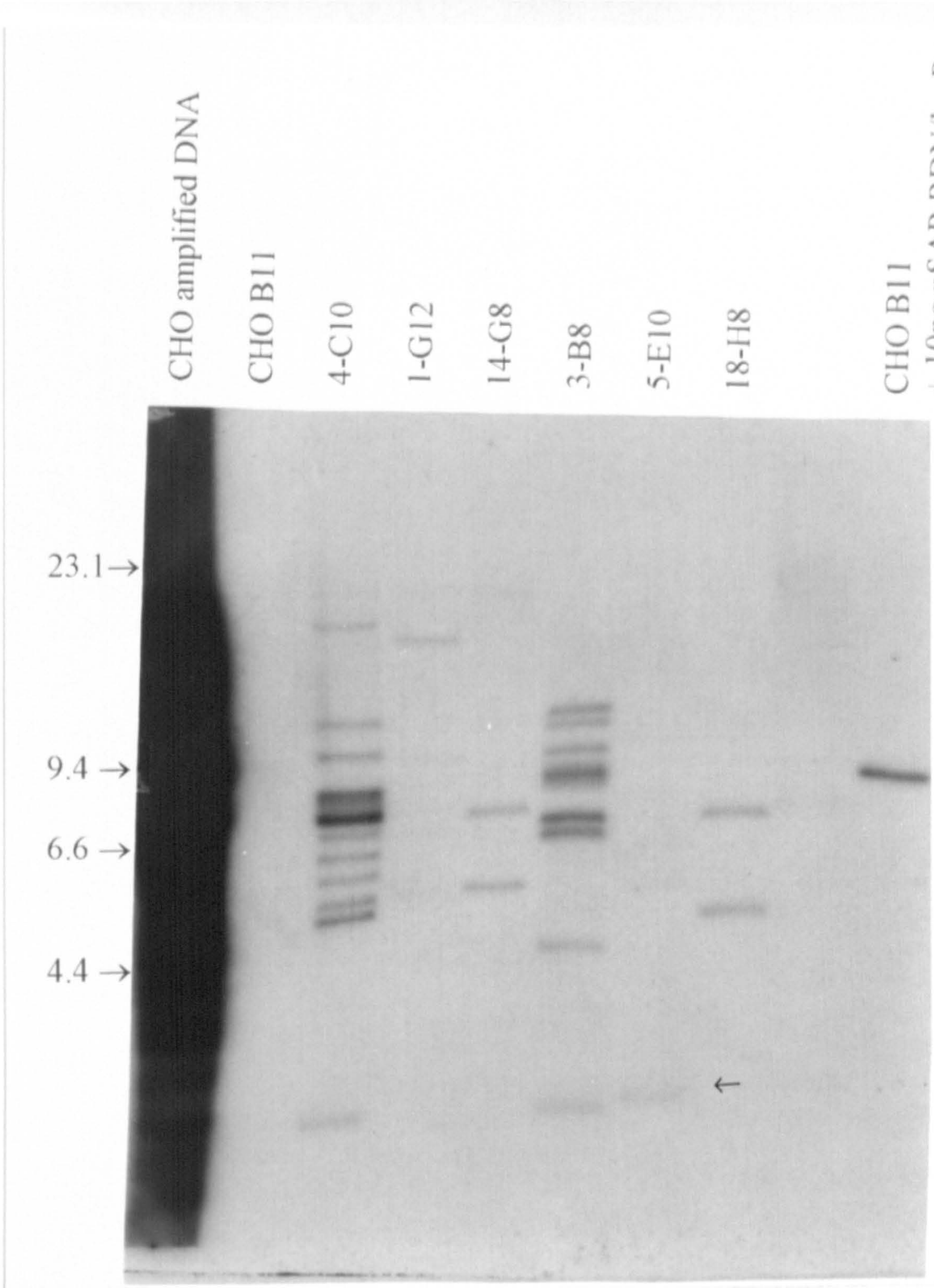


Figure 5.7a- Detection of pSAP-RDN/loxP (1 μg linearised) transfection into CHO-DUKX B11 cells by electroporation.

Genomic DNA (10 μg) prepared from independent CHO-DUKX B11 transformants was digested with *Bgl*II. The digested DNA was separated by electrophoresis on an agarose gel, and DNA fragments transferred onto a nylon membrane and probed with a radiolabelled SV40 polyA fragment. Lanes are labelled according to clone number. Clones 1-G12 and 5-E10 each shows a single signal. Clones 14-G8 and 18-H8 each shows two signals. The other clones shows multiple signals. The positive control (CHO-DUKX B11 + pSAP-RDNloxP) shows a signal of ~10 kb. DNA marker sizes to the left are in kilobases.

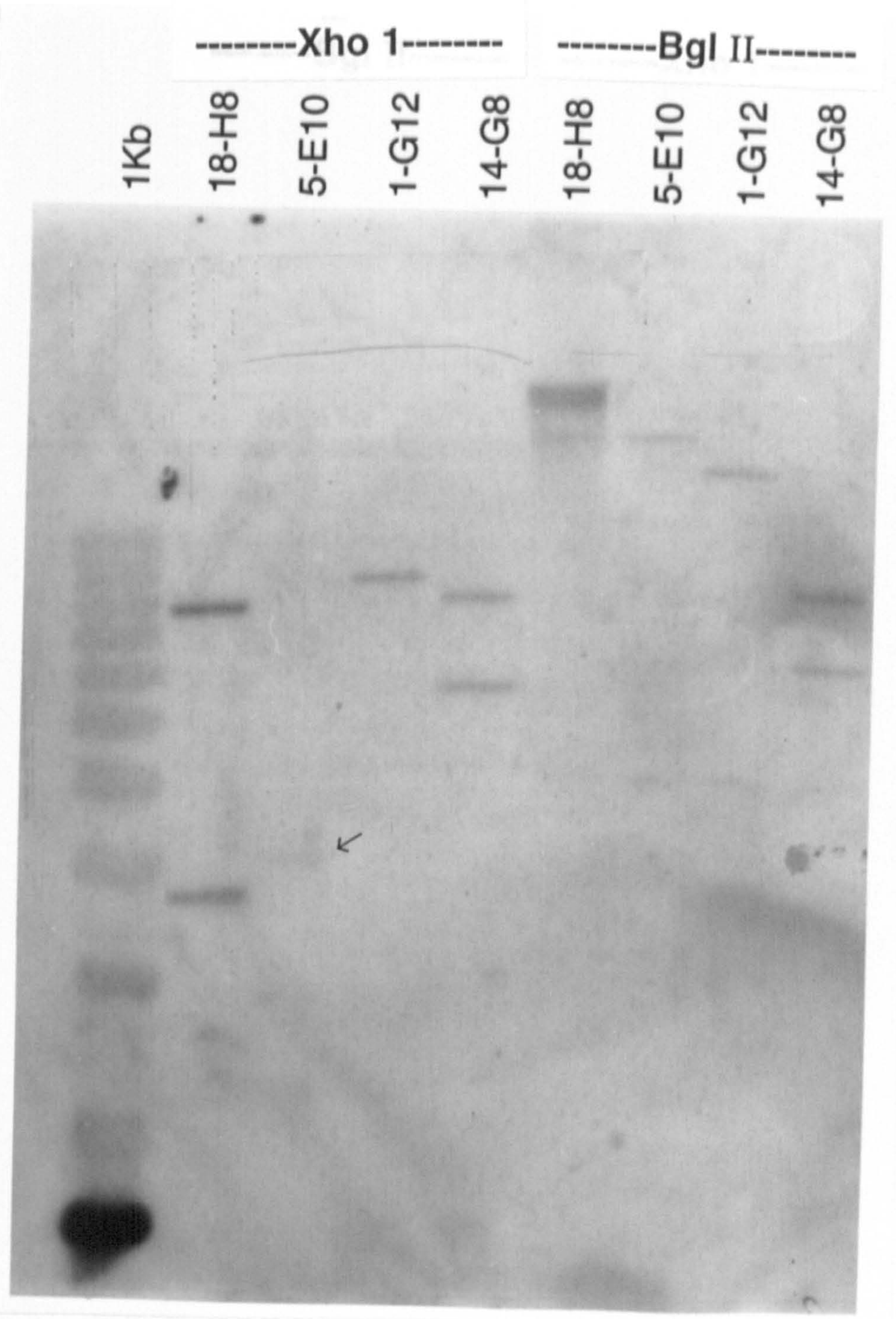


Figure 5.7b- Detection of transformed CHO-DUKX B11 cells containing either a single loxP plasmid or tandem repeats of the loxP plasmid by Southern blot analysis. Genomic DNA (10 μg) prepared from CHO B11 transformants was digested with either *Bgl*II or *Xho*l and analysed by Southern hybridisation as described in fig 5.7a. Lanes are labelled according to clone number. 1 Kb corresponds to standard molecular weight markers.

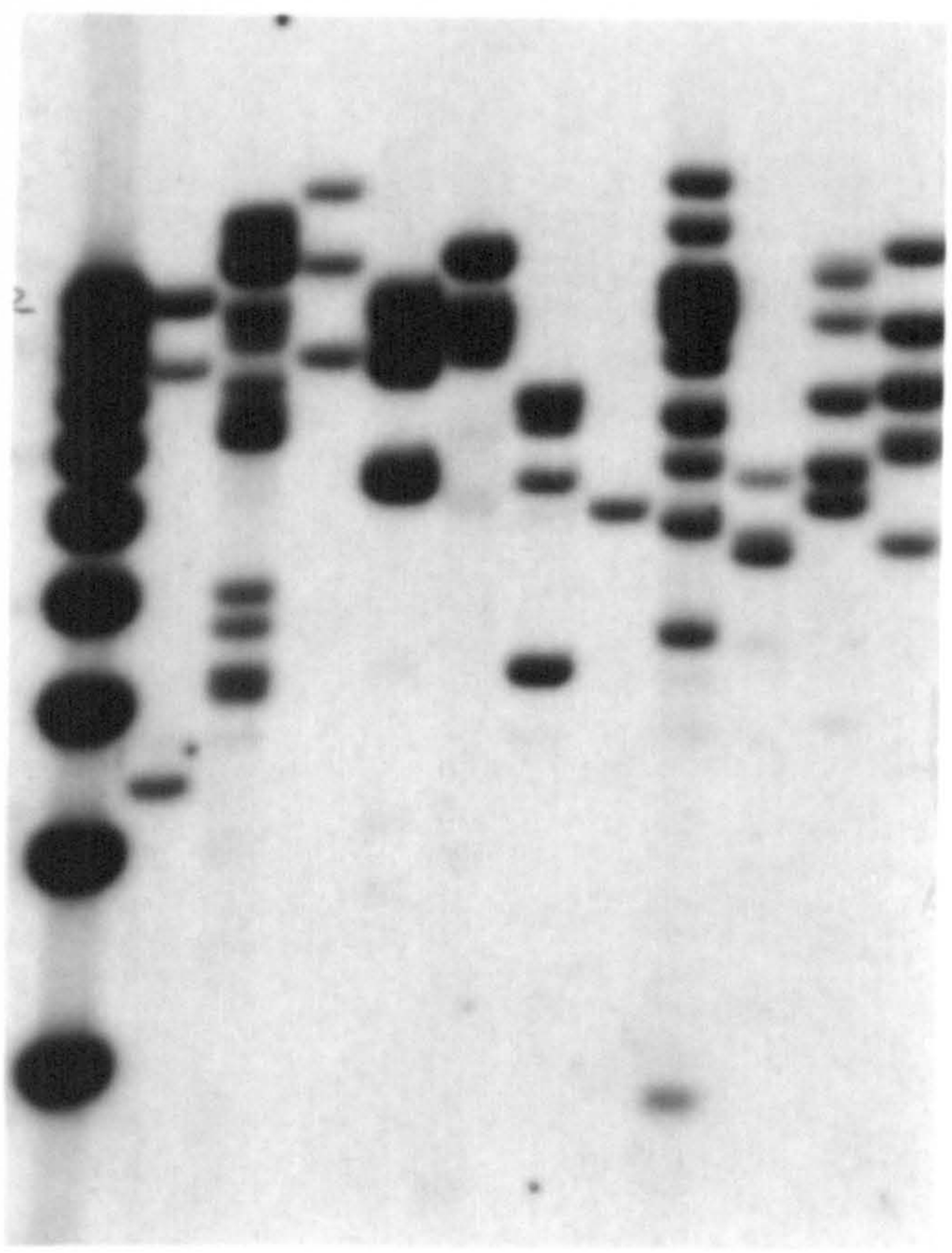


Figure 5.8- Detection of pSAP-RDN/loxP (1 μg linearised) transfection into CHO-DUKX B11 cells by electroporation.

Genomic DNA (10 μg) prepared from independent CHO-DUKX B11 transformants were digested with *Bgl*II. The digested DNA was separated by electrophoresis on an agarose gel, and DNA fragments transferred onto a nylon membrane and probed with a radiolabelled SV40 polyA fragment. Lanes are labelled according to clone number. 1 Kb corresponds to standard molecular weight markers.

signal could actually contain a single copy of the loxP insert and that the weak signal could be a partial digestion of genomic DNA. Unfortunately, clone 7-D1 was not investigated further because the cells got contaminated during the propagation steps.

Results from Southern analysis thus far shows that transfecting linearised loxL plasmid by electroporation leads to the isolation of stably transfected cell lines containing a single copy insertion of the loxL plasmid. Although, only three such cell lines were obtained, these cell lines can be used to investigate Cre-mediated site-specific integration events and to obtain information about chromosomal position effects on gene expression. To investigate chromosomal position effects on SAP expression, the specific production rate (SPR) of SAP was determined for the 3 cell lines each containing a single copy insertion of loxP site as described in the next section.

5.6.2.1 Investigating SPR of SAP in DHFR$^+$ clones containing a single copy insertion of loxL plasmid

Because each of the three cell lines contains a single copy of the loxL plasmid randomly integrated into the CHO chromosome, any variation in SAP expression should be due to different chromosomal position effect on gene expression. The specific production rate (SPR) of SAP was determined as described in section 2.8.4.

Results presented in Figure 5.9 shows that the three cell lines 13-G6, 1-G12 and 5-E10 showed a high, medium and a low level of SAP expression respectively. The high SAP expressing clone, 13-G6 showed a five fold higher expression compared to the medium SAP expressing clone, 1-G12 and a hundred fold higher expression compared to the low SAP expressing clone, 5-E10. This result suggests that compared to cell lines 5-E10 and 1-G12, cell line 13-G6 contains a loxL plasmid integrated into a chromosomal region more favourable for gene expression.

5.6.3 Summary of CHO transfected cells containing a single copy of loxP site

Three CHO transfected cell lines has been constructed each containing a single copy of loxP insert by transfecting 1 μg of linearised pSAP-RDN/loxP (loxL) into CHO DUKX B11 cells by electroporation. The three cell lines showed different levels of SAP expression. Because each cell line contains a single copy of the SAP gene (under

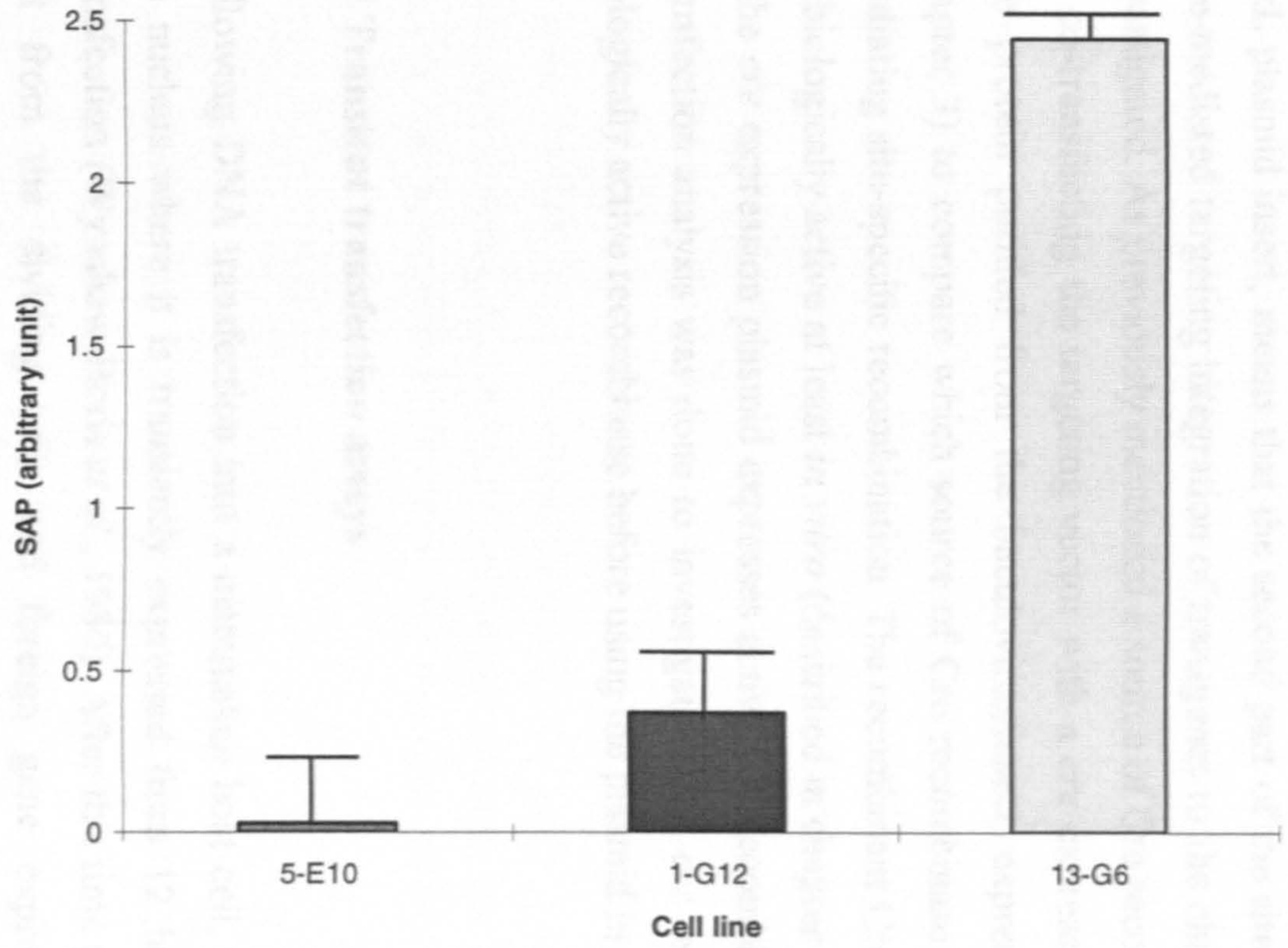


Figure 5.9- Comparison of the specific production rate (SPR) of SAP for three cell lines each containing a single copy of loxP site. A million cells were grown in 5 ml of $DHFR^+$ select medium overnight. The medium was then replaced with 5 ml of fresh medium. 24 hours after the addition of the fresh medium (+/- 30 min), 0.5 ml of medium was removed to determine SAP activity (described in section 2.4.7). The adherent cells in the flask were detached by versene and trypsin treatment and viable cell density determined. Arbitrary unit is defined as the change in absorbance at A_{405nm} in 30 minutes (readings were done in duplicates).

SPR (arbitrary unit/10^6 cells/24 hours)= $\frac{\text{Arbitrary unit of SAP activity}}{\text{cell density } (10^6/\text{ml})}$

the transcriptional control of a β-actin promoter) randomly integrated into the CHO chromosome, the different levels of SAP expression is a direct indication of different chromosomal position effect on SAP expression. Thus, cell line expressing a high SAP activity suggests that the SAP gene has integrated into a chromosomal region favourable for gene expression, whilst the cell line expressing low SAP activity suggests that the SAP gene has integrated into a chromosomal region less favourable for gene expression.

The identification of CHO transfected cell lines each containing a single copy of the loxL plasmid insert, means that the second part of the site-specific recombination (i.e. Cre-mediated targeting integration of transgenes to the chromosomal loxP site) can be investigated. As previously mentioned a source of Cre recombinase was to be provided by co-transfecting the targeting vector with a *cre* expression plasmid or recombinant Cre protein purified from the baculovirus/insect expression system (described in chapter 3) to compare which source of Cre recombinase would be more efficient at mediating site-specific recombination. The recombinant Cre protein has been shown to be biologically active at least *in vitro* (described in chapter 4), however it is not known if the *cre* expression plasmid expresses active Cre recombinase. Therefore, a transient transfection analysis was done to investigate if the *cre* expression plasmid expresses biologically active recombinase before using the plasmid in stable transfections.

5.7 Transient transfection assays

Following DNA transfection into a mammalian host cell, the DNA is translocated to the nucleus where it is transiently expressed from 12 hours up to 72 hours post-transfection (Wynshaw-Boris *et al.*, 1986). After this time plasmid copies are gradually lost from the dividing cells and foreign gene expression stops unless stable transfections, that contain a copy of the plasmid capable of replication either by the presence of a viral replicon in the vector or by integration into the host genome, are selected.

Transient transfections have a number of advantages over stable transfections: (i) levels of throughput especially in the study of large numbers of small-scale transfections e.g. comparison of different ratios of substrate plasmid to protein plasmid, (ii) they are rapid to analyse, with expression of foreign gene being assayed from 1-3 days post-infection. In contrast, stable transfections require approximately two to four weeks

selection in order to isolate stable transformants. (iii) the expression level of the foreign gene in the transfected vector is not influenced by the number of vector copies that integrate (except in COS cells were vectors containing the SV40 replicon and can replicate to very high copy number; $>10^4$ copies/cell), or by the position of integration into the host genome, as in the case of stable transfection (Kaufman *et al.*, 1985). (iv) potential DNA rearrangements that may occur during integration into the genome are avoided.

A limitation to transient transfection studies is the potential variation in transfection efficiency. Various strategies have been used to solve this problem including; the use of "internal control" plasmids (encoding reporter proteins), whose activity may be measured in each transfection. Comparison of the levels of the reporter protein activity in each transfection may indicate any variation in the transfection efficiency, which can be taken into account when analysing results by normalising the expression levels of the test construct to the reporter protein to be a reliable measure of transfection efficiency (Lake and Owen, 1991). A second method to control for variation in transient transfection efficiency, described by Abken and Reifenrath (1992), involves determining the amount of plasmid DNA transfected into the cells in each transfection, by "dot-blot" procedure. Despite these limitations transient transfection assays continues to be used for investigating various kinetic parameters of protein reactions *in vivo*. In this study, transient transfection assay was used to establish if *cre* expression plasmid encodes active Cre recombinase and also to gain an idea of the amount of substrate DNA and Cre plasmid DNA required for site-specific recombination to take place *in vivo*.

5.7.1 Investigating the activity of Cre expression plasmid in transient transfection assays

A simple way to assess if the Cre plasmid encodes active recombinase was to transfect the plasmid into cells to mediate intramolecular (excision) recombination event because a quantifiable assay system for the intramolecular recombination event has previously been described (section 3.7). Transfection was done using Transfectam (a synthetic cationic lipopolyamine molecule whose mode of action is described in section 5.4), instead of electroporation because the number of plasmid DNA copies transfected into cells was not essential whilst at the same time an efficient way of introducing DNA into cells was needed.

Initially CHO-DUKX B11 (2.5×10^5) cells were transfected with 20 μg of the *cre* plasmid, pMC1-Cre (using 60 μg of transfectam reagent) to determine if the *cre* plasmid expressed functional Cre recombinase in CHO cells. Crude extracts prepared from CHO transfected cells 24 hours post infection were used in an *in vitro* assay to investigate Cre-mediated recombination using plasmid pAT-lacZloxP2 as substrate. Results shown in Table 5.2 indicated that the crude extract prepared from CHO-DUKX B11 cells transfected with pMC1-Cre expressed a low level of active recombinase which can mediate site-specific recombination *in vitro*.

This initial observation suggested that the Cre plasmid could mediate intramolecular recombination *in vivo*. To investigate this, CHO-DUKX B11 (2.5×10^5) cells were transfected with 10 μg of pAT-lacZloxP2 and different amounts of the Cre expressing plasmid, pMC1-Cre using 60 μg of Transfectam for each transfection. In addition CHO-DUKX B11 (2×10^5) cells were co-transfected with 10 μg pAT-lacZ (negative control) and 20 μg pMC1-Cre, to investigate if Cre recombination is specific at loxP sites. In another experiment CHO-DUKX B11 (2×10^5) cells were co-transfected with 10 μg of pAT-lacZloxP2 and 20 μg of plasmid pSAP-RDN (which does not contain the *cre* gene) to investigate if the host recombinases could catalyse site-specific recombination at the loxP sites in pAT-lacZloxP2.

Plasmid DNA was prepared from transfected cells, by harvesting the cells either 24 or 48 hours post infection and Cre-mediated recombination analysed as described in section 3.7.2. Basically, aliquots of plasmid DNA (~ 50 ng) were transfected into *E. coli* DH5α competent cells and plated onto L-agar (supplemented with amp, X-gal and IPTG). The ratio of white to blue colonies observed on L-agar plate was taken to be the % recombination efficiency.

Results presented in Table 5.3 showed that Cre-mediated recombination occurred when 10 μg of the substrate plasmid (pAT-lacZloxP2) and 20 μg of Cre plasmid were used in the transfection. It is possible that increasing the amount of the Cre plasmid would increase the efficiency of the recombination. However, it is also possible that the presence of more Cre protein could catalyse the integration of the excised molecules reducing the efficiency of the excision event. The recombination efficiency was observed to be 18% 24 hour post-transfection and 6% 48 hours post-transfection. The difference in the recombination efficiency (measured as % excision) suggests that Cre recombinase activity was maximal before 48 hours post infection. Alternatively the substrate plasmid must have been degraded after 24 hour post infection, therefore

Source of recombinase	% Recombination
Crude cell extract (1 µl)	0
Crude cell extract (2 µl)	0
Crude cell extract (5 µl)	1.5
Crude cell extract (10 µl)	2
Untransfected crude cell extract (5 µl)	0

Table 5.2- Investigating Cre recombinase activity in crude cell extract.

CHO-DUKX B11 (2×10^5 cells) was transfected with 20 µg of pMC1-Cre. Different volumes of crude cell extract prepared from transfected cells (24 hours post transfection) were incubated with 2 µg of pAT-lacZloxP2 in an *in vitro* recombination assay. After termination of each reaction, aliquots (containing ~50 ng of DNA) were transfected into *E. coli* DH5α competent cells and dilutions analysed on L-agar plates (containing ampicillin, X-gal and IPTG). The ratio of white and blue colonies corresponding to recombinants and substrates respectively (see section 3.7.II) is a measure of the percentage (%) recombination.

Amount of Cre plasmid, pMC1-Cre (μg)	% Recombination (24 hours post transfection)	% Recombination (48 hours post transfection)
5	0	0
10	0	0
15	9	2
20	18	6

Table 5.3- Investigating the minimum ratio of substrate to Cre protein required for the *in vivo* site-specific recombination.

Plasmid pAT-lacZloxP2 (10 μg) was co-transfected with different amounts of the *cre* expression plasmid, pMC1-Cre in transient transfections assay using CHO-DUKX B11 cells as host. Plasmid DNA was prepared from transfected CHO-DUKX B11 cells either 24 or 48 hours after transfection. Aliquots (~ 100 ng) of the plasmid DNA were transfected into *E. coli* DH5α competent cells and plated onto L-agar (supplemented with ampicillin, X-gal and IPTG). The ratio of white to blue colonies observed on the L-agar plate was taken to be the % recombination efficiency.

reducing the amount of substrate DNA for Cre. However, the amount of substrate DNA at various times after the transfection was not determined using techniques such as "dot-blot". When recombination resulting from co-transfecting pAT-lacZ and *cre* plasmid into CHO cells were analysed on L-agar plate, only blue (17) colonies were observed. No white (recombinant products) colonies were observed indicating that the Cre protein is specific for the loxP site (data not shown). Furthermore, in the experiment, where pAT-lacZloxP2 was co-transfected with pSAP-RDN into CHO cells, recombination events analysed on L-agar plate produced only blue (9) colonies, no white colonies were observed suggesting that recombination at loxP sites is mediated by Cre protein and not host recombination protein(s) (data not shown). These results demonstrates that site-specific recombination in vivo is dependent on the presence of Cre recombinase and occurs specifically at loxP sites.

The observation that the *cre* plasmid expressed maximal Cre recombinase under 48 hours post transfection is in agreement with previous report by Sauer and Henderson (1988). These authors showed in the swine kidney cell line, PK15 that Cre mediated recombination activity accumulated to near its maximal level 15 hours post-infection by monitoring the generation of Cre-mediated excision reaction at various time points in an *in vivo* assay reaction. They also showed that there was a lag of about 5 hours in the Cre-mediated recombination compared to the accumulation of Cre protein. This contrasts the rapid Cre-mediated recombination *in vitro* which takes place within a few minutes (reported in chapter 4). The difference in the kinetics of Cre recombination *in vitro* and *in vivo* is due to a different cellular environment. For instance, during transfection the *cre* plasmid must be transported to the nucleus where the *cre* gene is transcribed and subsequently transported back into the cytoplasm where it is assembled into active Cre protein before recombination can take place. Furthermore, mammalian cells contain many DNA binding proteins which might be expected to compete with Cre protein for its loxP sites so perhaps preventing productive encounters between the Cre protein and the substrate DNA. This may explain the lag in Cre-mediated recombination compared to the accumulation of Cre protein.

The observations made from the transient transfections shows that the *cre* plasmid can produce active Cre protein to mediate site-specific targeting of transgenes to a chromosomal loxP site, however it should be borne in mind that the kinetics of the *in vitro* integration analysis observed in chapter 4, may be different from the integration event in CHO cells.

5.8 Site-specific targeting of transgenes to a chromosomal loxP site

As previously mentioned, CHO cell lines containing a single copy of loxP insertions were to be used as host expression systems to investigate Cre-mediated precise targeting of any exogenous DNA to the chromosomal loxP sites. For the integration to occur, the incoming DNA must also contain a loxP site. Therefore two different targeting vectors each containing a single loxP site were constructed to investigate the integration event.

5.8.1 Construction of loxP targeting plasmids

The two targeting vectors are pMAMneoCAT-loxP and pHM3-loxP. These vectors were designed to contain different reporter genes as indicators for the site-specific recombination event (described in detail in sections 5.7.2 and 5.7.3).

Plasmid pMAMneoCAT-loxP was constructed by cloning a loxP site in a promoterless pMAMneoCAT vector. Plasmid pMAMneoCAT (Clontech) contains the chloramphenicol acetyltransferase (CAT) locus as a reporter gene to provide a positive control for expression from an inducible MMTV-LTR promoter. The plasmid also contains the neomycin resistance gene, under the transcriptional control of the SV40 early promoter to permit positive selection of transformed cells, the β-lactamase gene conferring ampicillin resistance (Amp^r) and pBR322 ori for propagation of the plasmid in bacteria. To construct pMAMneoCAT-loxP (Figure 5.10), pMAMneoCAT was digested with *Nde*I and *Nhe*I to remove the entire MMTV-LTR promoter. The digestion gave rise to two fragments, a large fragment of about 7 kb comprising a promoterless CAT vector and a small fragment of about 2 kb comprising the MMTV-LTR promoter. The two fragments were separated on an agarose gel and the large fragment purified from the gel. This vector was dephosphorylated (to prevent self-ligation in a ligation reaction) and used in a ligation reaction with a double stranded oligonucleotide sequence comprising loxP site and the necessary protruding ends for ligation. The ligation products were analysed by transfecting an aliquot of the ligation reaction into *E. coli* DH5α competent cells and transformed cells plated onto L-agar supplement with ampicillin. Results showed that in the control ligation reaction (containing only the vector DNA) only one colony was observed on L-agar plates whereas in the ligation between the loxP sequence and the vector DNA thirteen colonies were observed suggesting that about 90% ligation between loxP and pMAMneoCAT. Sequence analysis was done to check that the recombinant DNA

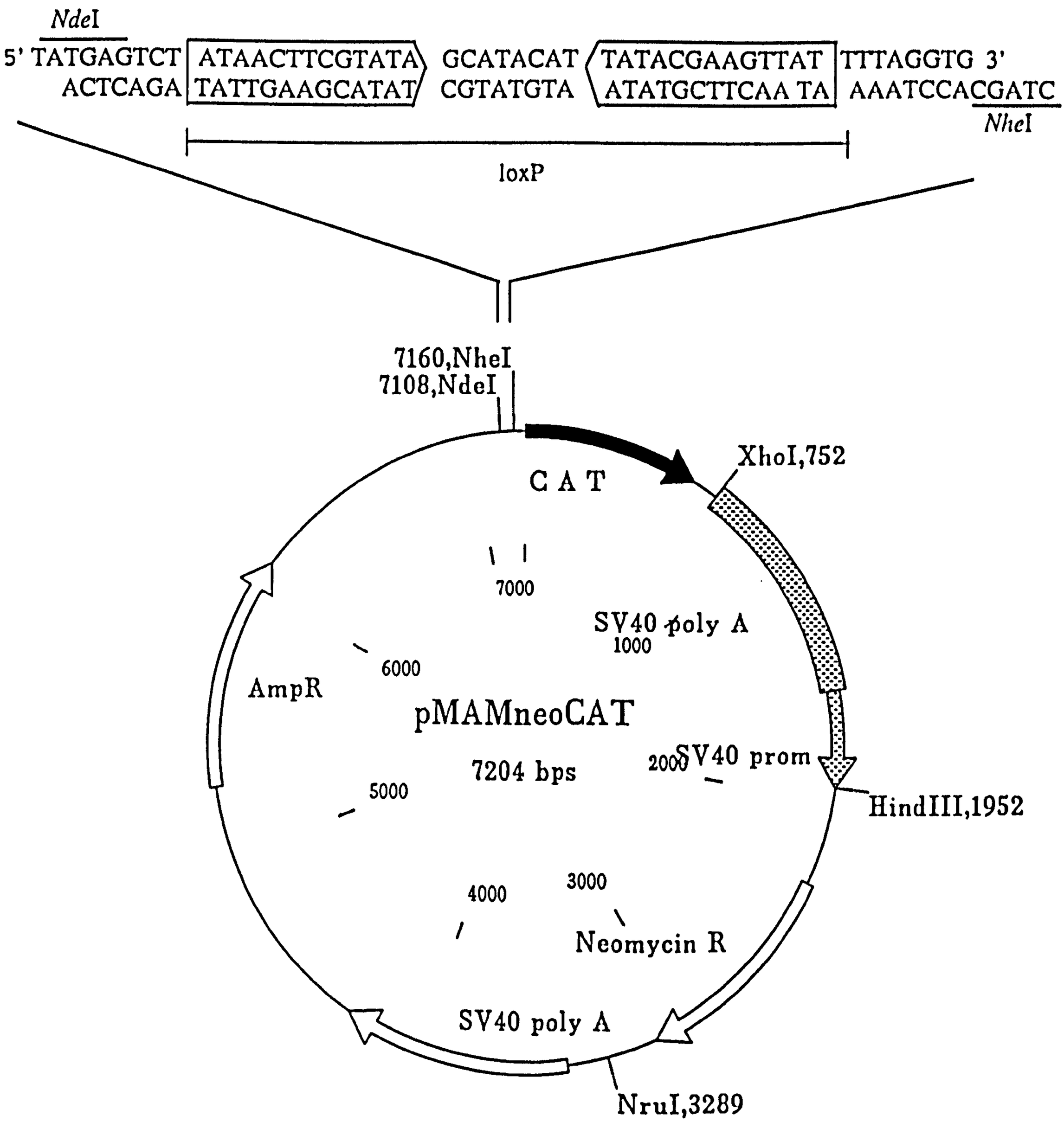


Figure 5.10- Map of pMAMneoCAT-loxP

A double stranded oligonucleotide sequence corresponding to the loxP site (which consists of two 13 bp inverted repeats [shown in box with arrowheads] flanking an 8 bp spacer region [shown in bold]) with *Nde*I and *Nhe*I recognition ends was cloned into pMAMneoCAT digested with *Nde*I and *Nhe*I restriction endonucleases.

contained the loxP insert, and in the right orientation and that the sequence corresponds to the loxP site (data not shown).

Plasmid pHM3-loxP (Figure 5.11) was constructed by cloning a loxP site into pHM3. Plasmid pHM3 (Kaestner *et al.*, 1994) contains a promoterless *lacZ* coding region of *E. coli* (reviewed in Alan and Cook, 1990) and neomycin resistance (neo^r) marker (Southern and Berg, 1982) fused to generate a single open reading frame (ORF) followed by the poly(A) of mouse phosphoglycerate kinase (*pgk*) gene. A multiple cloning site and the polyadenylation signal of the human growth hormone (GH)-encoding gene (to terminate any spurious transcripts that might arise from cryptic promoters within the vector) is present upstream of the *lacZ*/neo^r fusion gene. The plasmid also contains the β-lactamase gene (*bla*) and ori from plasmid pUC19. The fusion gene has been shown to be bi-functional in embryonic stem (ES) cell chimaeras (Zachgo and Gossler, 1993) and is especially useful when a promoterless gene targeting construct is to be used in homologous recombination experiments to decrease the frequency of illegitimate recombination and to increase the chance of obtaining a homologous recombinant clone (reviewed in Porter, 1989).

To construct pHM3-loxP, a double stranded oligonucleotide comprising the loxP sequence with the neccessary protruding ends for ligation was cloned into the unique *Hind*III site in pHM3. In order to bias for the ligation of the loxP sequence and the vector, the vector was dephosphorylated (see section 2.2.9) after digestion with the *Hind*III restriction endonuclease. Results of the ligation event was analysed by transfecting ligated DNA fragments into *E. coli* DH5α competent cells and transformed cells observed as ampicillin resistant colonies on L-agar plate containing ampicillin (see section 2.2.13). The results showed forty-four colonies from the control plate (corresponding to self-ligated vector DNA) whilst only fifteen colonies were observed on the test plate (i.e. ligation between pHM3 and the loxP sequence). Because of the high proportion of self-ligated vector DNA, DNA was prepared from cultured samples of all the colonies observed on the test plate and digested with *Pvu*II. Digestion of pHM3 DNA with *Pvu*II gives rise to five fragments of approximately 2.9, 2.8, 0.5, 0.3, and 0.22 kb sizes. If however, loxP sequence in inserted into *Hind*III site in pHM3, then *Pvu*II digestion is expected to increase the 0.5 kb fragment by the 50 bp loxP fragment which can be observed on an agarose gel. Results of PvuII digestions showed that three out of the fifteen colonies analysed had the loxP sequence insert (data not shown). DNA from these three colonies were sequenced to further confirm that they contain the loxP sequence and also to check that the sequence corresponded to loxP (data not shown).

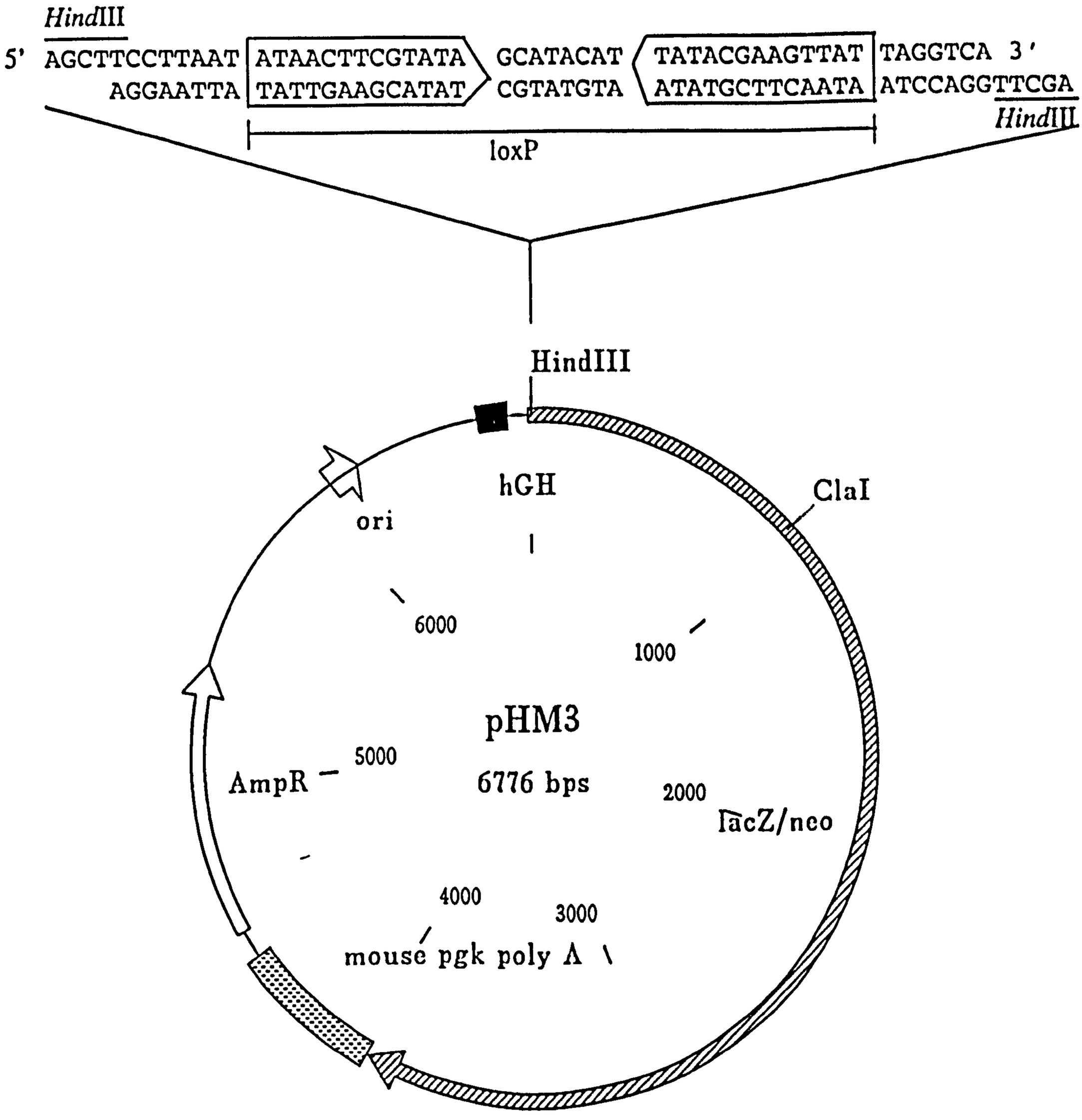


Figure 5.11- Map of pHM3-loxP.

A double stranded oligonucleotide sequence corresponding to the loxP site (which consists of two 13 bp inverted repeats [shown in box with arrowheads] flanking an 8 bp spacer region [shown in bold]) with the *Hind*III recognition ends was cloned into the unique *Hind*III site in pHM3.

The two plasmids, i.e. pMAMneoCAT-loxP and pHM3-loxP were independently used as targeting vectors to investigate Cre-mediated site-specific integration into chromosomal loxP site in cell line 1-G12 (which showed the medium SAP expression level out of the three cell lines identified to contain a single copy of the loxP insert; see section 5.5.2.1).

5.8.2 Cre expression plasmid mediated site-specific targeting of pMAMneoCAT-loxP into cell line 1-G12.

To investigate if Cre recombinase can direct the precise integration of an exogenous DNA (containing a loxP site) to a chromosomal loxP site, pMAMneoCAT-loxP was co-transfected with the *cre* expression plasmid, pMC1-Cre into cell line 1-G12.

The targeting vector was designed such that Cre-mediated integration of the promoterless CAT gene to the chromosomal loxP site would integrate the CAT gene at the loxP site which is downstream from the β-actin promoter (Figure 5.12). The result of site-specific recombination should therefore activate the expression of CAT gene from the β-actin promoter. In addition, because transcription of the SAP reporter gene was under the control of the β-actin promoter, site-specific recombination should therefore separate the SAP gene from the β-actin promoter. Thus the recombination event should not only give rise to the expression of CAT activity (which can be quantitatively assayed by an *in vitro* ELISA [section 2.8.2]) but also the loss or reduction of SAP activity.

Cell line 1-G12 (1×10^7 cells) was transfected with 10 μg of pMAMneoCAT-loxP and 20 μg of either pMC1-Cre or pBS (Bluescript plasmid; as a negative control for the targeting event) by electroporation at 960μF and 250V. This ratio of the amount of substrate to *cre* plasmid was used because this same ratio was used in transient transfections to show site-specific recombination. Two days after transfection, transfected cells were split three fold and grown in $DHFR^+$ select medium supplemented with G418 to 400 μg/ml final concentration. Adherent transfected cells were re-fed after 6 days to remove non-transfected cells and cell debris.

The transfection efficiency was approximately found to be 1.5 times higher when pMAMneoCAT-loxP was co-transfected with pBS than when co-transfected with pMC1-Cre; approximately 1100 colonies/10^7 transfected cells compared to 740 colonies/10^7 cells (calculated from crystal violet staining of individual resistant

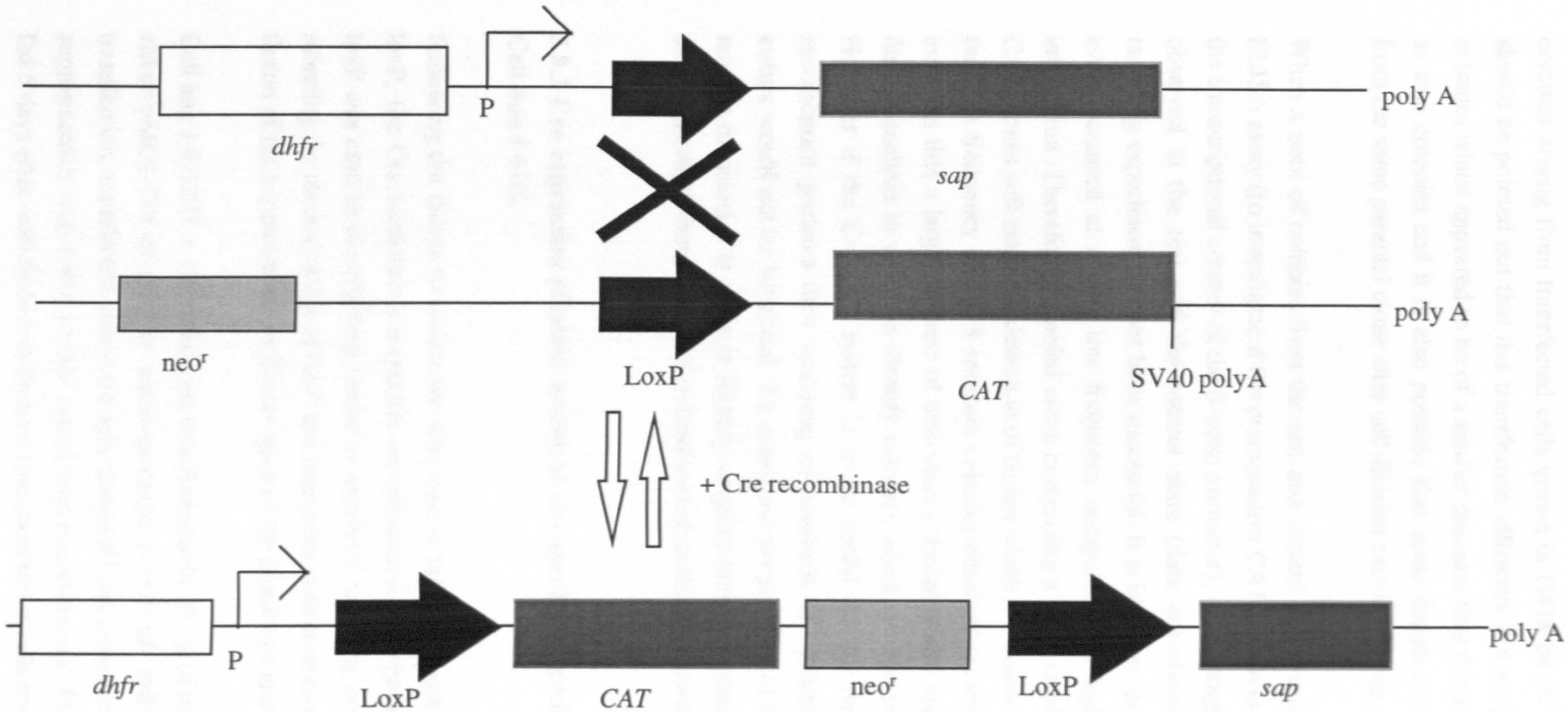


Figure 5.12-Schematic representation of Cre-mediated targeting of an exogenous DNA to a chromosomal loxP site.
Initially a loxP site was placed in the genome of CHO DHFR$^-$ cells via a plasmid encoding a dihydrofolate reductase gene (*dhfr*) for selection and a secreted alkaline phosphatase (*sap*) reporter gene under the transcriptional control of β–actin promoter (P). Cre recombinase mediated targeting of a promoterless *CAT* gene to the chromosomal loxP site should activate the expression of the CAT reporter gene from the β-actin promoter initially driving the SAP expression.

colonies arising from transfected cells grown in G418 by 20 days post infection). It should be pointed out that this transfection efficiency is a rough estimate, since a few colonies which appeared to be of a smaller diameter than the average were disregarded as true colonies and it is also possible that some daughter colonies which originate from the same parental clone after cell division may have been counted twice.

When a pool of colonies from the test and control plates were used in a quick CAT ELISA assay (to investigate if the promoterless CAT gene was being transcribed under the transcriptional control of the β-actin promoter), only background CAT activity was observed in the test and the control plate (data not shown), suggesting that the targeting experiment had not been successful. It is however, possible that the targeting event occurred at a very low frequency compared to a high frequency of random integration. Therefore a pooled clone comprising a large number of "non-expressing" CAT clones will mask the detection of clones which are expressing CAT. Furthermore, the high frequency of G418 resistant colonies observed on the control and test plates indicates that a large numbers of individual colonies would need to be propagated for further analysis in order to identify colonies which are the result of targeting event. However if the Cre-loxP system is to be useful for the large scale expression of recombinant proteins then screening of hundreds of colonies to identify targeted events would not be beneficial. To overcome the problem of having to screen a large number of colonies in order to identify targeted events, plasmid pHM3-loxP was used as the targeting vector for the Cre-mediated recombination event.

5.8.3 Cre expression plasmid mediated site-specific targeting of pHM3-loxP into Cell line 1-G12

Following the failure to isolate the site-specific targeted event using pMAMneoCAT-loxP, the Cre mediated site-specific recombination was repeated. This time, pHM3-loxP was used as the targeting vector to enrich for targeting over illegitimate events by selecting for the activation of *lacZ* and neomycin resistance as a result of Cre-mediated fusion of the promoterless *lacZ*/neo^r gene to the β-actin promoter.

Cell line 1-G12 (1×10^7 cells) was transfected with 10 μg of pHM3-loxP and 20 μg of either pMC1-Cre or pBS by electroporation at 960 μF and 250V. Two days after transfection, transfected cells were split three fold and grown in DHFR$^+$ select medium supplemented with G418 to 400 μg/ml final concentration. Transfected cells were re-fed 6 days after transfection to remove non-transfected cells and cell debris.

The transfection efficiency was approximately 1.6 fold higher when pHM3-loxP was co-transfected with pBS than when co-transfected with pMC1-Cre; ~240 colonies/10^7 transfected cells compared to ~150 colonies/10^7 transfected cells (calculated from crystal violet staining of individual resistant colonies arising from transfected cells resistant to G418 by 20 days post infection). The observation of a high number of G418 resistant colonies on the control plate suggested that random integration also leads to the activation of the neo^r gene. Because the *neo*^{r} gene is fused to the *lacZ* gene (encoding β-galactosidase), it is possible that the *lacZ* gene is also activated. An advantage for using plasmid pHM3-loxP to monitor Cre-mediated recombination event is that a large number of colonies can be rapidly screened in situ for β-gal expression. Therefore, β-gal staining was used to investigate the frequency of *lacZ* activation as a result of illegitimate recombination and compared to that as a result of Cre-mediated targeted integration.

5.8.3.1 β-gal staining of G418 resistant clones

To analyse G418 resistant colonies for β-galactosidase expression, a third of the colonies isolated from the control transfection (~ 80 colonies on a large petri dish) and a third of the colonies from the test plate (the Cre-mediated reaction; ~ 50 colonies) were stained for β-gal activity (Lim and Chae, 1989; see section 2.8.5 for the staining reaction). G418 resistant colonies which expressed β-galactosidase were observed under a light microscope as blue colonies (data not shown). Of the 50 colonies resulting from Cre-mediated reaction, 13 were positive for β-gal staining. Whilst only 3 out of 80 colonies from the control transfection were positive for β-gal staining. This result correlates to approximately 26% of the colonies from the Cre-mediated reaction expressing β-galactosidase whilst only about 4% of the colonies resulting from random integration activated β-galactosidase expression. Because the targeting vector contains a promoterless *lacZ* and *neo*^{r} coding sequences fused to generate a single open reading frame, the high frequency of G418 resistant colonies expressing β-galactosidase from the Cre-mediated reaction suggests that the activation of both the *lacZ* and *neo*^{r} is the result of Cre-mediated targeted integration to the β-actin promoter. In contrast, the low frequency of G418 resistant colonies which expressed β-galactosidase demonstrates that spontaneous activation of both *lacZ* and *neo*^{r} as a result of illegitimate recombination occurs at low frequency.

As a result of this initial observation that the Cre plasmid can mediate site-specific integration of pHM3-loxP vector to a chromosomal loxP site, G418 resistant colonies resulting from the targeted and non-targeted events were propagated for further analysis.

5.8.3.2 β-gal and SAP assays of targeted clones

As previously mentioned, the result of the targeting event should not only activate the expression of *lacZ*/neo^r fusion gene but also give loss or reduction of SAP activity. In order to investigate this, 15 G418 resistant clones from Cre-mediated event and 6 clones from the control reaction were assayed for β-gal activity (see section 2.8.6 for method) and SAP activity and compared to the parental cell line (Table 5.4). Results of β-gal assay showed that the average clone isolated from the Cre-mediated reaction expressed about 3.5 fold higher β-galactosidase than the control clones. Although the level of *lacZ* expression from a β-promoter was not investigated (because the only way of determining the expression level of a single copy insertion of the *lacZ* reporter gene to a specific chromosomal position would require Cre-mediated site-specific recombination), the observation of higher β-galactosidase expression from the Cre-mediated reaction compared to random integration suggests that spontaneous activation of *lacZ*/neo^r gene occurs at chromosomal regions other than the β-actin promoter locus.

The average level of SAP activity from the targeted clones was less than half compared to the parental, indicating that the SAP phenotype is not completely abolished by the targeting event. This result suggests a residual SAP expression in the targeted clones which could be the result of transcriptional read-through from the β-actin promoter. In contrast, the average SAP expression level from the control clones with the exception of one clone was about the same as the parental cell line. The low expression level of SAP observed in this clone was due to the poor growth rate of cells in this clone.

The results from β-galactosidase and SAP assays suggests that both the Cre-mediated targeting and random integration events leads to the activation of β-galactosidase expression and a reduction in SAP activity. Although there are differences in the expression levels of β-galactosidase and SAP from the targeted and the non-targeted

Transformant		β-galactosidase (mU)	SAP (arbitrary unit/10^6 cells/day)
1-G12 (parental)		-	0.37
Targeted	1	1.2	0.1
	2	3.3	0
	3	2.5	0.17
	4	1.8	0.08
	5	2.9	0.3
	6	2.3	0.24
	7	1.9	0.4
	8	2.5	0.12
	9	1.7	0.08
	10	3.4	0.1
	11	1.9	0
	12	3.1	0.09
Mean		**2.37**	**0.15**
Random	13	0.2	0.26
	14	0.8	0.3
	15	1.9	0.25
	16	0.7	0.39
	17	0.4	0.16
	18	-	0.34
Mean		**0.66**	**0.28**

Table 5.4- Estimation of β-galactosidase and SAP (SPR) activity from independent pHM3-loxP transfected 1-G12 cells.

One million cells from (cell line 1-G12) transformants were assayed for β-galactosidase activity (described in section 2.4.10) and SAP activity (described in section 2.4.8).Results of β-galactosidase and SAP activity are the average of assays from duplicate extracts from each transformant. Background β-galactosidase activity in 1-G12 parent cell line was subtracted from activities of the transformants. One unit of β-galactosidase hydrolyses 1 micromole of o-nitrophenyl-β-D-galactopyranoside (ONPG) to o-nitrophenol and galactose per minute at pH 7.5 at 37°C. Arbitrary unit of SAP is defined as the change in absorbance at A_{405nm} in 30 minutes (readings were done in duplicates).

events, these differences do not prove site-specific recombination. Therefore, further investigations of the targeting event was necessary.

5.8.3.3 PCR analysis of targeted clones

A PCR was used to investigate if Cre-mediated recombination juxtaposed the pHM3-loxP to the chromosomal loxP site. A combination of oligonucleotide primer sequences designed as forward primer to anneal within β-actin promoter (CB1736-GCTATTCTCGCAGGATCAGTCGACC) and a reverse primer (CB1739-ATAATTCGCGTCTGGCCTTCCTGTA) to anneal within and the *lacZ* gene were used to amplify sequences from genomic DNA prepared from clones isolated from the Cre-mediated reaction (Figure 5.13). If pHM3-loxP is targeted to the chromosomal loxP site, PCR analysis using these primers is expected to produce a 590 bp fragment .

Results of the PCR reaction showed a lot of bands in all of the targeted clones analysed including a 590 bp fragment band (indicated by arrow; lanes 4-12) which probably corresponds to the targeted event. Unfortunately, a very faint band migrating at about the same distance as the 590 bp was also observed in the control reaction using genomic DNA from the parental cell line 1-G12 (lane 2). In addition to this faint band, other bands were observed in the control reaction which migrated to the same distance as those seen in the clones isolated from the Cre-mediated reaction. This suggests that the 590 bp fragment like the other bands are the result of the primer sequences amplifying genomic DNA sequences non-specifically. Various PCR parameters such as Mg^{2+} salt concentration and the annealing temperature were varied in the hope of improving the specificity of the primer sequences. Furthermore, different primers were designed and used to investigate the targeting event. Unfortunately, none of the PCR reactions done unambiguously showed the expected DNA fragment to suggest targeted event. Consequently, a Southern hybridisation was done to investigate the targeting event.

5.8.3.4 Southern blot analysis of targeted clones

Genomic DNA (10 μg) prepared from the clones isolated from the Cre-mediated reaction and control clones was digested with *Bgl*II and the digested DNA separated by gel electrophoresis on an agarose gel. The DNA fragments were transferred onto a

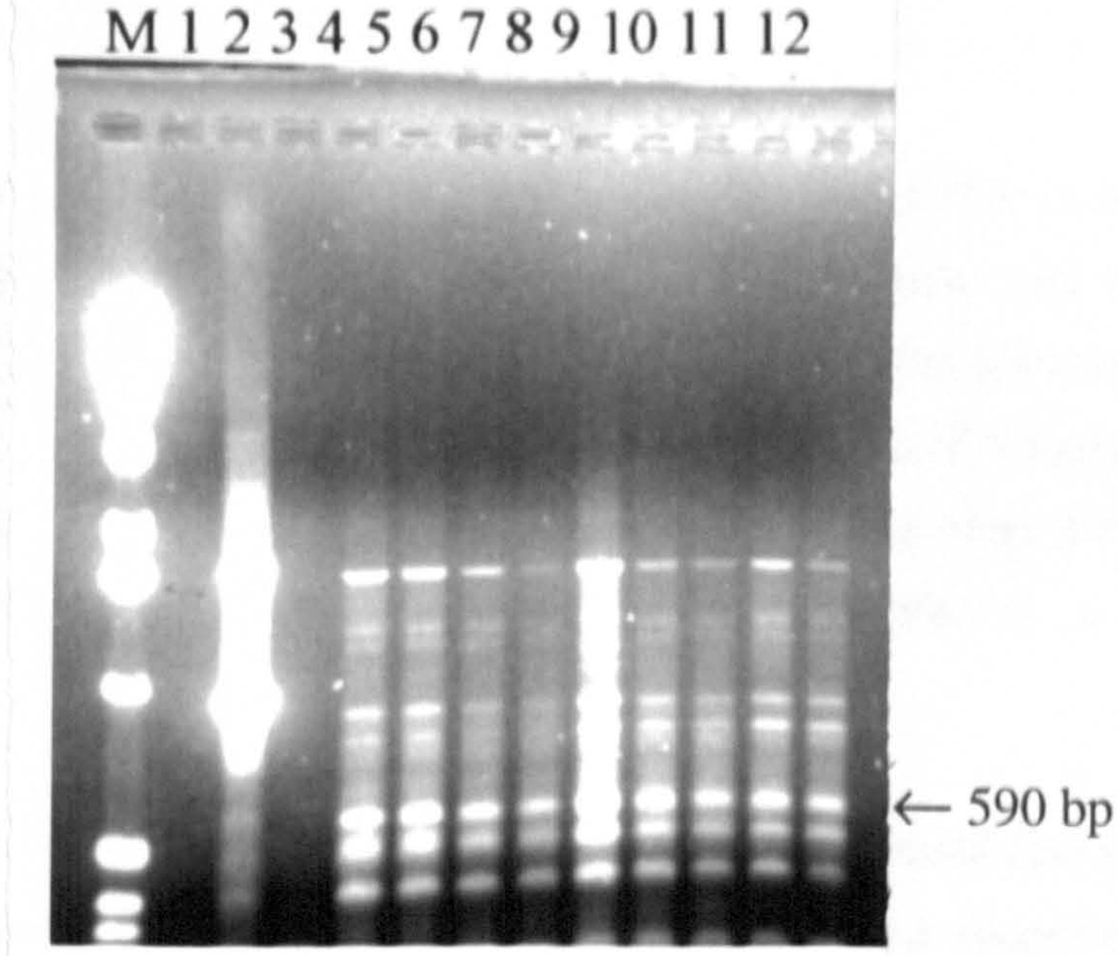


Figure 5.13- Agarose gel separation of PCR reaction of independent clones resulting from pHM3-loxP tranfection into cell line 1-G12.

PCR reactions were done using genomic DNA prepared from independent cell line 1-G12 transformants. A pair of oligonucleotide primers designed to prime within the β-actin UTR region and the *lac*Z gene was used to amplify sequence from the genomic DNA. M indicates the 1 Kb molecular weight DNA marker; lane 1; negative control reaction (primer sequence only), lane 2; genomic DNA from 1-G12 (negative control), lane 3; genomic DNA from CHO-DUKX B11 (negative control), lanes 4-12; genomic DNA from independent clones representing Cre-mediated transfection of pHM3-loxP into cell line 1-G12. A 590 bp fragment band is indicated to the right of the gel picture.

nylon membrane and hybridised with a radiolabelled *lacZ* gene (980 bp fragment obtained by digesting pHM3 with *Cla*I and *Hind*III).

The restriction enzyme *Bgl*II will cut the genomic DNA in many places, however, if Cre-mediated recombination had occurred efficiently and integrated the targeting vector, pHM3-loxP to the chromosomal loxP site, then genomic digestion with *Bgl*II is expected to produce a 7.7 kb fragment. When the lacZ fragment is used as a probe in the Southern blot, the DNA fragment containing the targeted fragment is expected to hybridise and appear on the autoradiograph as a signal of a 7.7 kb fragment (Figure 5.14a).

Results of Southern hybridisation (Figure 5.14b) showed that eight out of twelve G418 resistant and β-gal$^+$ clones (from the Cre-mediated reaction) analysed showed one signal (lanes 1, 2, 4, 8-12) which suggests single-copy integrants. However the signals were approximately 6.5 kb in size (1.2 kb less than would be anticipated for the targeted event). Three other targeted clones (lanes 5, 6 and 7) showed this 6.5 kb signal and other signals to suggest multiple insertions. The other targeted clones (lanes 3 and 14) gave different size signals characteristic of random integration or multiple insertions. The clones isolated from the control reaction (lanes 15-20) gave different signals to suggest random integration. However, the 6.5 kb signal was observed in lane 19 and appeared very faint in in the other control clones investigated. Because this 6.5 kb signal appears to be present in the control clones as well, this would suggests that the targeted event had not occurred. However, this signal is smaller than anticipated for the targeted event which could suggest a partial deletion of the targeting vector. If this 6.5 kb signal is indeed due to a partial deletion in the vector, then it is possible to observe this signal in both the Cre-mediated and the control events especially if it corresponds to a fragment of vector DNA. Furthermore, eight out of twelve clones from the Cre reaction only showed this 6.5 kb signal whereas the control clones showed multiple signals suggesting that this single copy integration is probably the result of Cre-mediated site-specific recombination. However, because of the size discrepancy between the observed and expected signals it was necessary to investigate the targeting event further.

To investigate the structural features of the single-copy integrants in more detail, other restriction digestion analyses were done. Genomic DNA from three targeted clones (single copy integrants) were digested with either *Hind*III, *Pvu*I or *Asp*718 + *Nru*I. Digested DNAs were separated on an agarose gel and DNA fragments transferred to a nylon membrane and hybridised with a radiolabelled *lacZ* probe as described above.

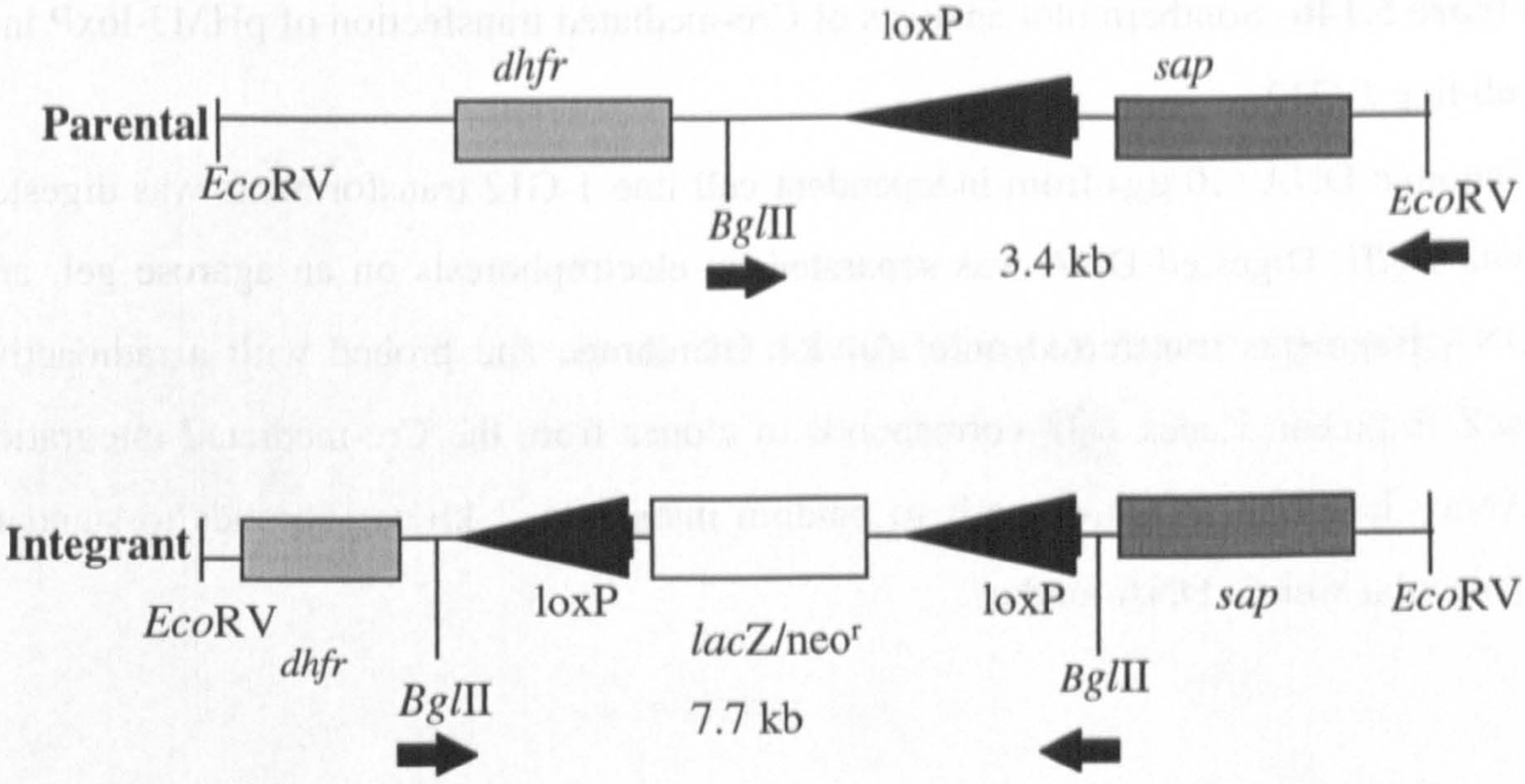
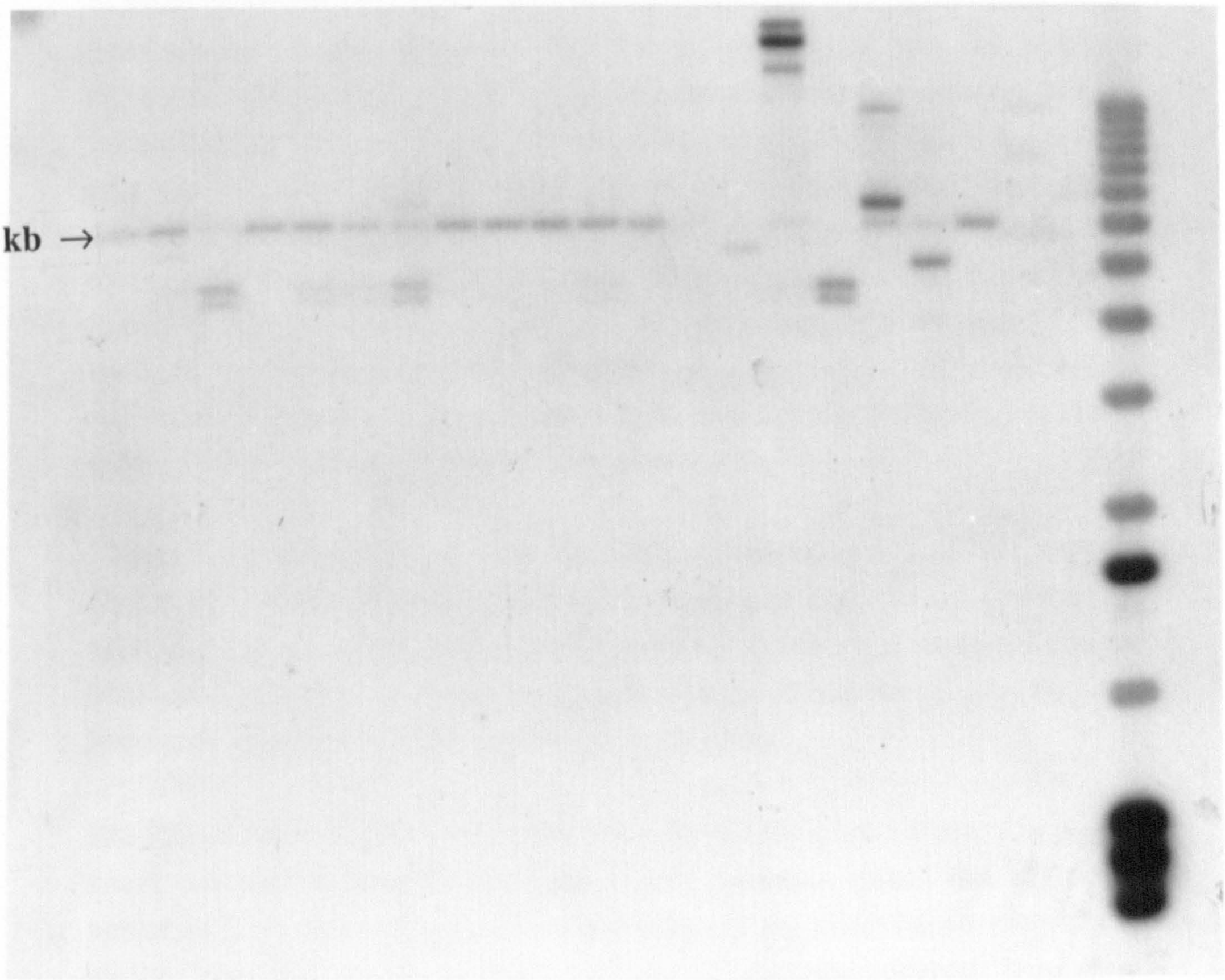


Figure 5.14a- Linearised map of the chromosomal target before Cre-mediated integration (parental)and after targeting with pHM3loxP.

Figure 5.14b- Southern blot analysis of Cre-mediated transfection of pHM3-loxP into cell line 1-G12.

Genomic DNA (10 μg) from independent cell line 1-G12 transformants was digested with *Bgl*II. Digested DNA was separated by electrophoresis on an agarose gel, and DNA fragments transferred onto a nylon membrane and probed with a radioactive *lac*Z fragment. Lanes 1-14 corresponds to clones from the Cre-mediated integration events; lanes 15-20 corresponds to random integrants. 1 kb corresponds to standard molecular weight DNA marker.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 1Kb
6.5 kb →

The Cre-mediated proper targeting event is expected to produce 6.7 kb; 17 kb plus unpredictable amount of genomic DNA; and 14 kb signals corresponding respectively to *Hind*III, *Pvu*I or *Asp*718 + *Nru*I restriction digestions in the Southern hybridisation (fig 5.15). The result showed an approximately 6 kb signal from the *Hind*III restriction digestion, a signal of more than 12 kb from the *Pvu*I restriction digestion and a signal of about 10 kb from *Asp*718 + *Nru*I restriction digestions. This signal pattern of the Southern hybridisation correlates with the expected sizes from the individual digestions, however the signals again migrated at smaller sizes than expected. This is a further indication that perhaps the targeting event had occurred but there is a deletion of a segment of the targeting vector probably by host cell endonucleases. Partial deletion of a targeting vector in a Cre-mediated recombination event has also been observed by Fukushige and Sauer, (1990). However, they reported a much lower frequency of vector deletion than observed here. To investigate if the anomaly of the predicted to the observed DNA fragment sizes is due to deletion would require restriction endonuclease digestion, and/or PCR, and sequence analysis of subcloned genomic DNA fragments prepared from targeted clones.

Although the clones isolated from the Cre-mediated reaction were not analysed further, the isolation of clones which were: G418 resistant, positive for β-galactosidase expression and showed a reduction in SAP activity (compared to parental) and also the observation of single copy integrants in the Cre-mediated reaction (all gave the same size signal) suggests that the targeted event had occurred..

The Cre-mediated targeting experiment was repeated using recombinant Cre protein which was purified from the baculovirus-insect expression system (see chapter 3) instead of a *cre* expression plasmid to investigate if the direct introduction of Cre protein into cells would be more efficient at mediating targeted integration. Unfortunately, cell line 1-G12, when revived from frozen (see section 2.3.5), grew very poorly therefore another cell line, 13-G6 which also contains a single loxP insertion was used as the host to repeat the targeted event. Cell line 13-G6 showed a higher expression level of SAP compared to 1-G12 suggesting that the chromosomal loxP site is in a region more favourable for gene expression. Therefore proper targeting of exogenous DNA to the loxP site in this cell line should give a characteristic higher level of β-galactosidase expression (compared to β-galactosidase expression from 1-G12) if the expression level of the targeted transgenes correlate with the SAP activity in the different cell lines. Thus proper targeting would also give information about the effect of different chromosomal regions on gene expression.

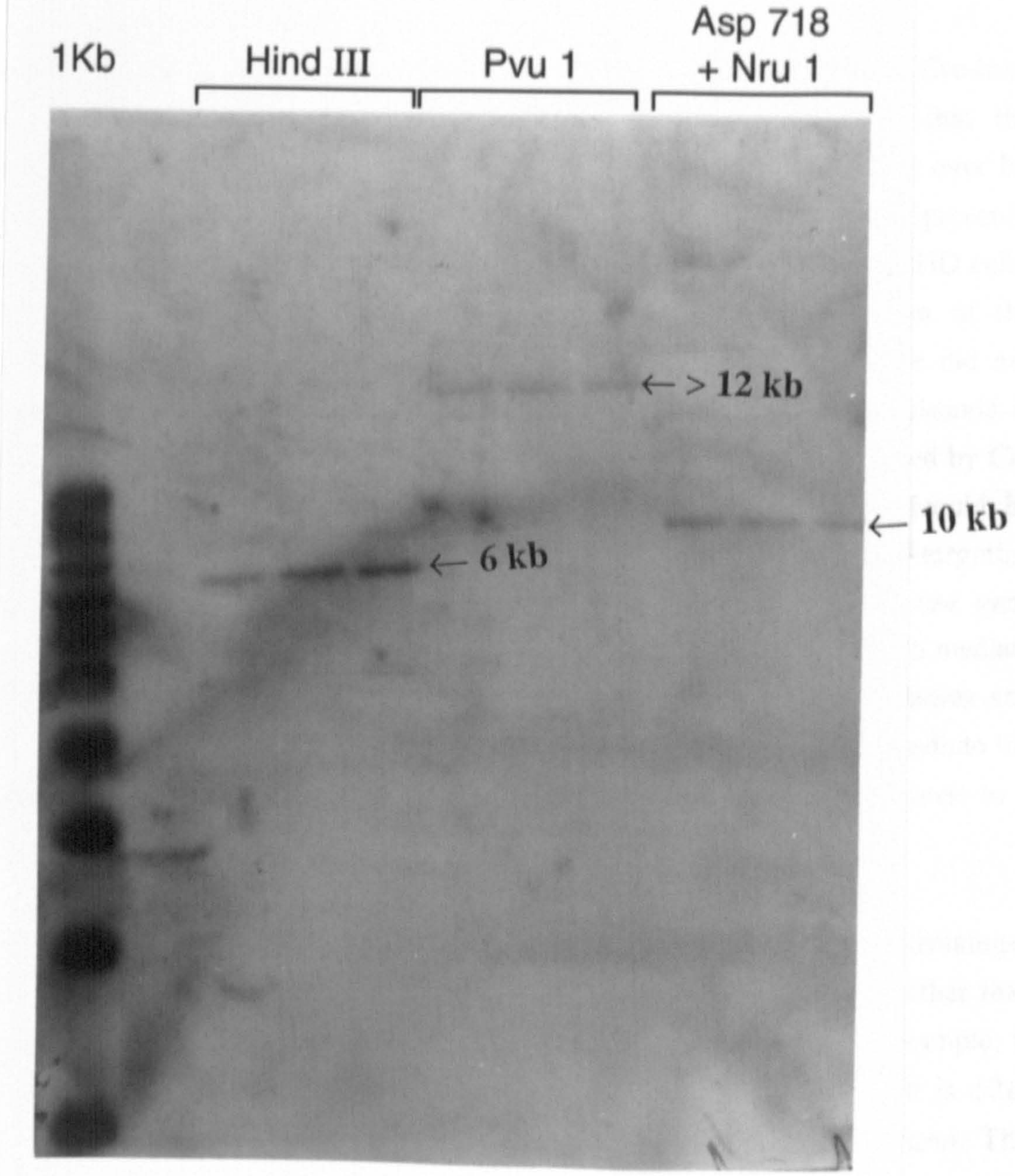


Figure 5.15- Detection by Southern blot analysis of single-copy integration of pHM3-loxP into 1-G12.

Genomic DNA (10 μg) from three independent Cre-mediated targeted transformants was each digested with either *Hind*III, *Pvu*I or *Asp*718 + *Nru*I. Digested DNA was analysed by Southern hybridisation using *lacZ* probe (essentially as described in Figure 5.14b).

5.9 Use of recombinant Cre protein in site-specific recombination

As previously mentioned in the introduction chapter, a drawback of the Cre-loxP system for targeted integration of transgenes into mammalian cells is that the recombination event is freely reversible, and excision is kinetically favoured over bi-molecular integration thus integrated products are inherently unstable in the presence of the recombinase. Therefore if the *cre* gene was stably integrated into the CHO cells, constitutive expression of the Cre recombinase will catalyse the excision of the integrated DNA. Southern analysis using radiolabelled *cre* gene as a probe did not show any signal (data not shown) to indicate the co-integration of the *cre* plasmid in the targeted event. However, any residual Cre activity in the cell (as evidenced by Cre activity in cells 48 hours post infection in transient transfection assay) would be expected to catalyse the re-excision of the integrated plasmid reducing targeting efficiency. A strategy that has been shown to obviate the possibility of *cre* gene integrating into cells was to directly introduce purified Cre protein into cells to mediate the targeting event (Baubonis and Sauer, 1993). Using this approach, Baubonis and Sauer showed that Cre protein purified from an *E. coli* source was able to mediate the integration of neomycin-phosphotransferase and thymidine kinase marker genes to a previously integrated loxP site in human osteosarcoma cell line, 143B.

The direct introduction of Cre protein into cells has a number of distinct advantages compared to using *cre* expression plasmids. First, use of purified protein, rather than DNA, eliminates the possibility of stable integration of the *cre* gene. For example, in the construction of targeted knockout mutations in embryonic stem cells it is often necessary to remove the selectable marker gene used to generate the knockout. This can be done by flanking the marker gene by directly repeated loxP sites in the targeting construct and later excising the marker gene in the resulting knockout cell line by transient expression of the *cre* gene (Sauer, 1993). Use of purified Cre protein eliminates the need to check marker gene "popout" colonies for the presence of the *cre* gene. Second because direct protein transfer should be possible with many different transfer cell types, there is no need for customised *cre* vectors having particular cell and tissue specific transcription and translation signals. Third, recombinase activity is temporally prescribed by the actual presence of functional Cre protein in the cell, thus averting problems for efficient integration of loxP targeting vectors arising from long term persistence (and expression) of DNA, as occurs in certain cells, such as striatal muscle (Wolfe *et al.*, 1992).

The recombinant Cre protein purified from baculovirus/insect expression system has been well characterised in two *in vitro* assay systems (chapter 4): First, gel retardation assays showed that the Cre protein was very specific for its recognition loxP site. Second, the Cre protein was shown to mediate intramolecular (excision) and intramolecular recombination in DNA substrates containing loxP sites. The *in vitro* characteristics of the Cre protein is identical to native Cre suggesting that it can be used to mediate site-specific recombination *in vivo*.

5.9.1 Cre protein mediated targeted integration into cell line 13-G6 using electroporation

The direct introduction of proteins into cells by electroporation (which occurs by the passive transport through temporary created pores in plasma membrane) has been shown for various proteins including viral transactivator (Kashanchi *et al.*, 1992) and antibodies (Berglund and Sarkey, 1991) and therefore should be possible for the recombinant Cre protein. Furthermore, the recombinant Cre protein contains an engineered nuclear localisation signal which should increase the rate of transport into the nucleus of the cell to mediate site-specific recombination.

Cell line 13-G6 (1×10^7 cells) was transfected with 10μg of pHM3-loxP and different amounts of recombinant Cre protein (i.e., 1, 5 and 10 μg; 27000 U/mg [section 3.8]) by electroporation at 960 μF and 250V. In a control experiment cell line 13-G6 (1×10^7 cells) was transfected with 10 μg of pHM3-loxP and 10 μg of pBS by electroporation at 960 μF and 250V. Two days after transfection, transfected cells were split three fold and grown in $DHFR^+$ select medium supplemented with G418 to 400 μg/ml final concentration. Transfected cells were refed 6 days after transfection to remove non-transfected cells and cell debris.

The transfection efficiency (calculated from crystal violet staining of individual colonies resistant to G418 by 20 days post infection) was similar between the control and the targeted events; ~175 colonies/10^7 transfected cells were obtained using 1 μg of Cre protein in the transfection; ~206 colonies/10^7 transfected cells were obtained using 5 μg of Cre protein in the transfection; ~193 colonies/10^7 transfected cells were obtained using 10 μg of Cre protein in the transfection whilst ~219 colonies/10^7 transfected cells were obtained using 10 μg of pBS in the transfection .

The observation of similar transfection efficiency between the control and the targeted events suggests that the recombinant Cre protein had not mediated very efficient targeting of pHM3-loxP to the chromosomal loxP sites. Because the recombinant Cre protein is clearly active, as shown by *in vitro* studies, the result suggests that perhaps the amounts of Cre protein investigated were not sufficient to observe site-specific recombination or that perhaps the electroporation conditions used (basically mixing the targeting vector and Cre protein on ice prior to transfection) were not optimum. Further investigations would require titrating different amounts of the Cre protein and the targeting vector and varying electroporation conditions. An alternative method that has successfully been used to investigate Cre-protein mediated recombination was lipofection (Baubonis and Sauer, 1993) therefore the targeting experiment was repeated this time transfecting the targeting vector and the recombinant Cre protein into cell line 13-G6 by lipofection.

5.9.2 Cre protein mediated targeted integration into cell line 13-G6 using lipofectACE

The mode of action of lipofectACE reagent is the same as transfectam reagent. Basically lipofectACE reagent interacts with DNA to form lipid-DNA complex which then associates with the cell plasma membrane. Internalisation of the DNA then occurs by endocytosis whilst internalisation of Cre protein should occur by passive transport into the cell because of its small size. The lipofectACE procedure requires transfections to be done in a six-well plate using about 10^5 cells. Because a small number of cells was to be used and also lipofection is a very efficient way of transfecting DNA into cells (as shown previously in section 5.4) 1 μg of the targeting vector was used in this transfection instead of the 10 μg used in transfection by electroporation.

Cell line 13-G6 (1×10^5 cells) was transfected with a constant amount of pHM3-loxP (1 μg) and either 1, 5 or 10 μg of recombinant Cre protein using lipofectACE (GIBCO BRL; see section 2.4.4 for procedure) to investigate optimal targeting (highest frequency of G418 resistance). In the control reaction, cell line 13-G6 (1×10^5 cells) was transfected with 1 μg of pHM3-loxP and 10 μg of pBS (Stratagene) using lipofectACE. Transfected cells were incubated in complete medium overnight. Twenty four hours post transfection, cells were fed with $DHFR^+$ select medium supplemented with G418 to 400 μg/ml final concentration. Cells were re-fed every six days to remove dead cells and cell debris.

Results showed 7 colonies/10^5 transfected cells using 1 µg of Cre protein; 27 colonies/10^5 transfected cells using 5 µg of Cre protein; 57 colonies/10^5 transfected cells using 10 µg of Cre protein and 22 colonies/10^5 cells using 10 µg of pBS (calculated from crystal violet staining of individual resistant colonies arising from transfected cells resistant to G418 by 20 days post infection). This is however an estimate on the assumption that there are no more colonies appearing 20 days. Results showed the highest targeting when 10 µg of Cre protein was used, this may not be the optimum amount of Cre required for targeting because higher amounts of Cre protein were not investigated.

Forty $DHFR^+$ and G418 resistant colonies were isolated from the targeted transfection using 10 µg of recombinant Cre protein and propagated for further analysis. Initially a Southern hybridisation analysis was used to investigate if Cre protein had directed the targeting of pHM3-loxP to the chromosomal loxP site (instead of by PCR because previous PCR analysis of the Cre-mediated event [see section 5.8.3.3] showed a lot of non-specific bands).

5.9.2.1 Southern hybridisation of the targeting event

Genomic DNA (10 µg) prepared from clones isolated from the Cre-mediated reaction and control clones was digested with *Bgl*II and the digested DNA separated by gel electrophoresis on an agarose gel. The DNA fragments were transferred to a nylon membrane and hybridised with a radiolabelled *lac*Z gene (980 bp fragment obtained by digesting pHM3 with *Cla*I and *Hind*III).

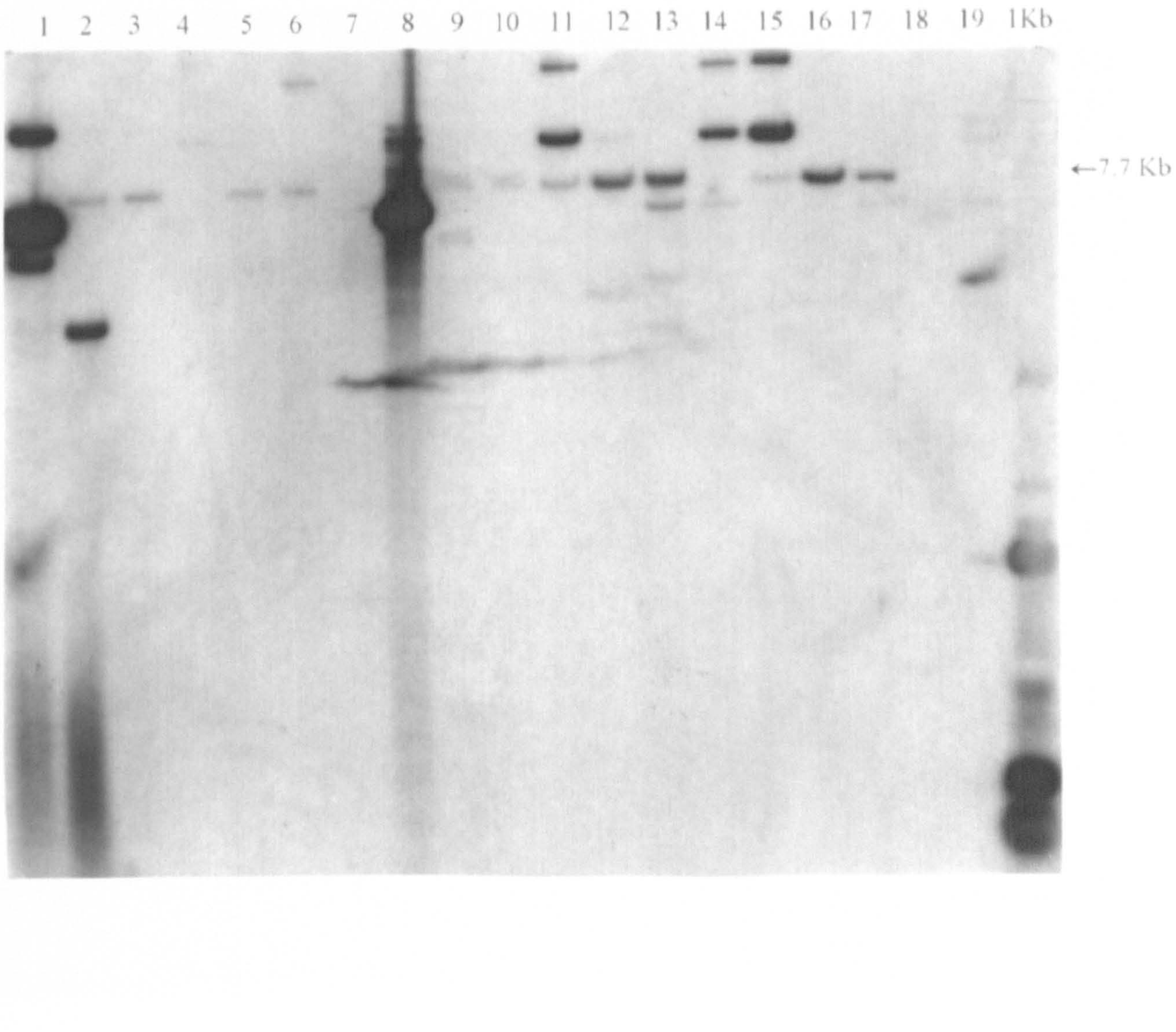
The restriction enzyme *Bgl*II will cut the genomic DNA in many places, however, if Cre-mediated recombination had occurred efficiently and integrated the targeting vector, pHM3-loxP to the chromosomal loxP site, then genomic digestion with *Bgl*II is expected to produce a 7.7 kb fragment. When the *lac*Z fragment is used as a probe in the Southern blot, the DNA fragment containing the targeted fragment is expected to hybridise and appear on the autoradiograph as a signal of 7.7 kb fragment (Figure 5.16).

Results of Southern hybridisation showed that out of 36 clones analysed 11 clones showed a single corresponding to a 7.7 kb fragment (lanes 3, 5, 10, 16 and 17 in Figure A; lanes 21, 22, 28, 29, 31 and 34 in Figure B) suggesting that about 30% of

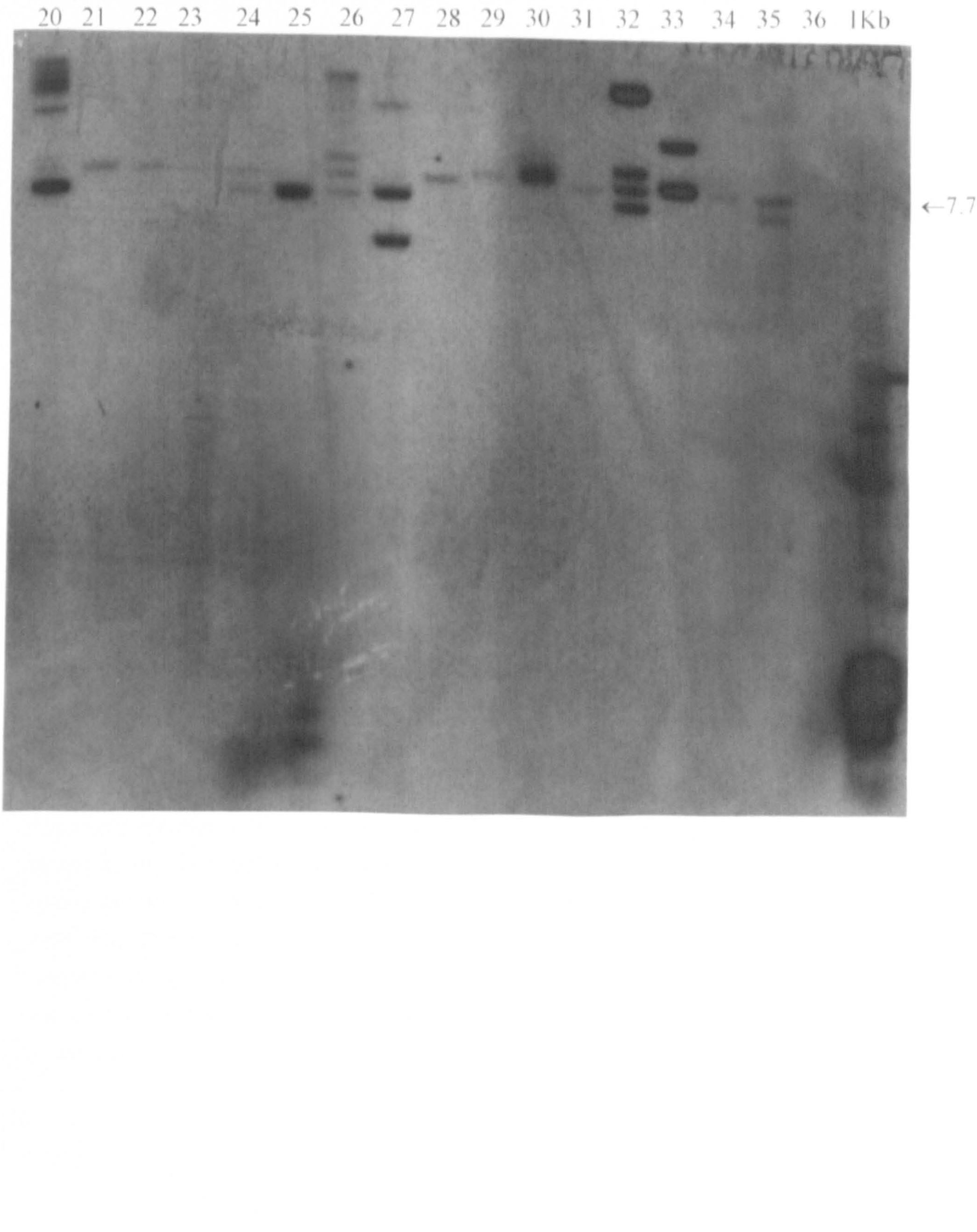
Figure 5.16- Detection of recombinant Cre protein mediated integration of pHM3-loxP into cell line 13-G6.

Genomic DNA (10 µg) prepared from independent transformants was digested with *Bgl*II. Digested DNA was separated by electrophoresis on an agarose gel, and DNA fragments transferred onto a nylon membrane and probed with a radioactive *lac*Z fragment. **A**- shows targeted clone numbers 1 to 19; **B**- shows targeted clone numbers 20 to 36. The 7.7 Kb indicated on the right corresponds to the expected size from a Cre-mediated integration event.

A



B



the clones analysed are single copy targeted integrants. However, some clones showed the 7.7 kb signal and other signals (lanes 2, 6, 11, 12, 13, 15 in Figure 5.16a; lanes 24; 25; 32 and 34 in Figure 5.16b) indicating targeted integrations as well as random. Other clones showed different signals indicating random integration or re-arrangements. To verify that the 7.7 kb signal observed as single copy integrant in some clones corresponds to targeted integration products, another Southern hybridisation was done.

5.9.2.2 Further Southern hybridisation to analyse the targeting event

Genomic DNA (10 μg) prepared from 10 clones which appeared to be single copy targeted integrants (observed as 7.7 kb signals, section 5.8.1) was digested with *Nru*I and *Cla*I. Digested DNA was separated on an agarose gel and DNA fragments transferred to a nylon membrane and probed with a radiolabelled 824 bp DNA fragment from pSAP-RDN (Figure 5.3) comprising a region of β-actin promoter (obtained by digesting pSAP-RDN with *Xho*I and *Bst*EII). Restriction digestion of the parental cell line 13-G6 with *Nru*I and *Cla*I is expected to produce a 6.8 kb fragment, however, if the Cre protein mediated proper targeting of pHM3-loxP to the chromosomal loxP site, then restriction digestion of genomic DNA from targeted clones with *Nru*I and *Cla*I is expected to yield a 5.9 kb fragment (Figure 5.17a). When the radiolabelled β-actin fragment is used as a probe in the Southern analysis, the DNA fragment containing the targeted fragment is expected to hybridise and appear on an autoradiograph as a signal of 5.9 kb whilst DNA from parental cell line would appear as a signal of 6.8 kb (Figure 5.17b).

Results showed the expected signal of 6.8 kb fragment in the parental clone. Seven out of the 10 clones investigated showed the expected 5.9 kb signal confirming that they are single integrants. Interestingly, these clones also showed the 6.8 kb parental signal. The 3 other clones (i.e transformants 5, 10, and 29) investigated showed signals which were smaller than the expected 5.9 kb fragment suggesting that they are not true single copy targeted integrants. However, the smaller signals could be due to a deletion in the targeting vector sequence which was not observed in the initial Southern analysis (using genomic DNA digested with *Bgl*II) because it occurred in sequences outside the *Bgl*II recognition sites. The observation of the signal corresponding to the parental fragment in the targeted clones shows that the targeted clones comprise a mixture of unrecombined parental cells and cells which have undergone the site-specific recombination event. Mixed colonies would be expected if

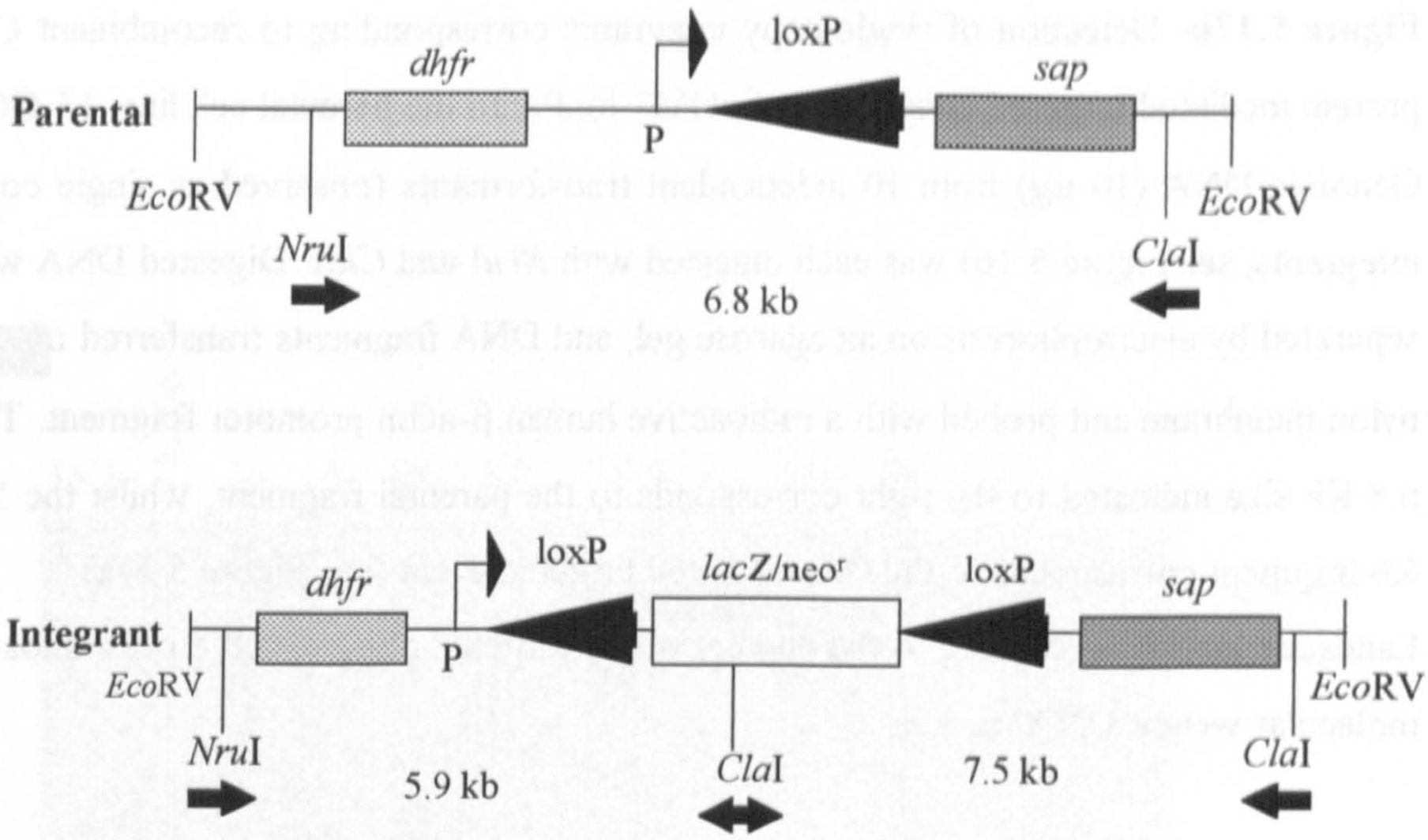
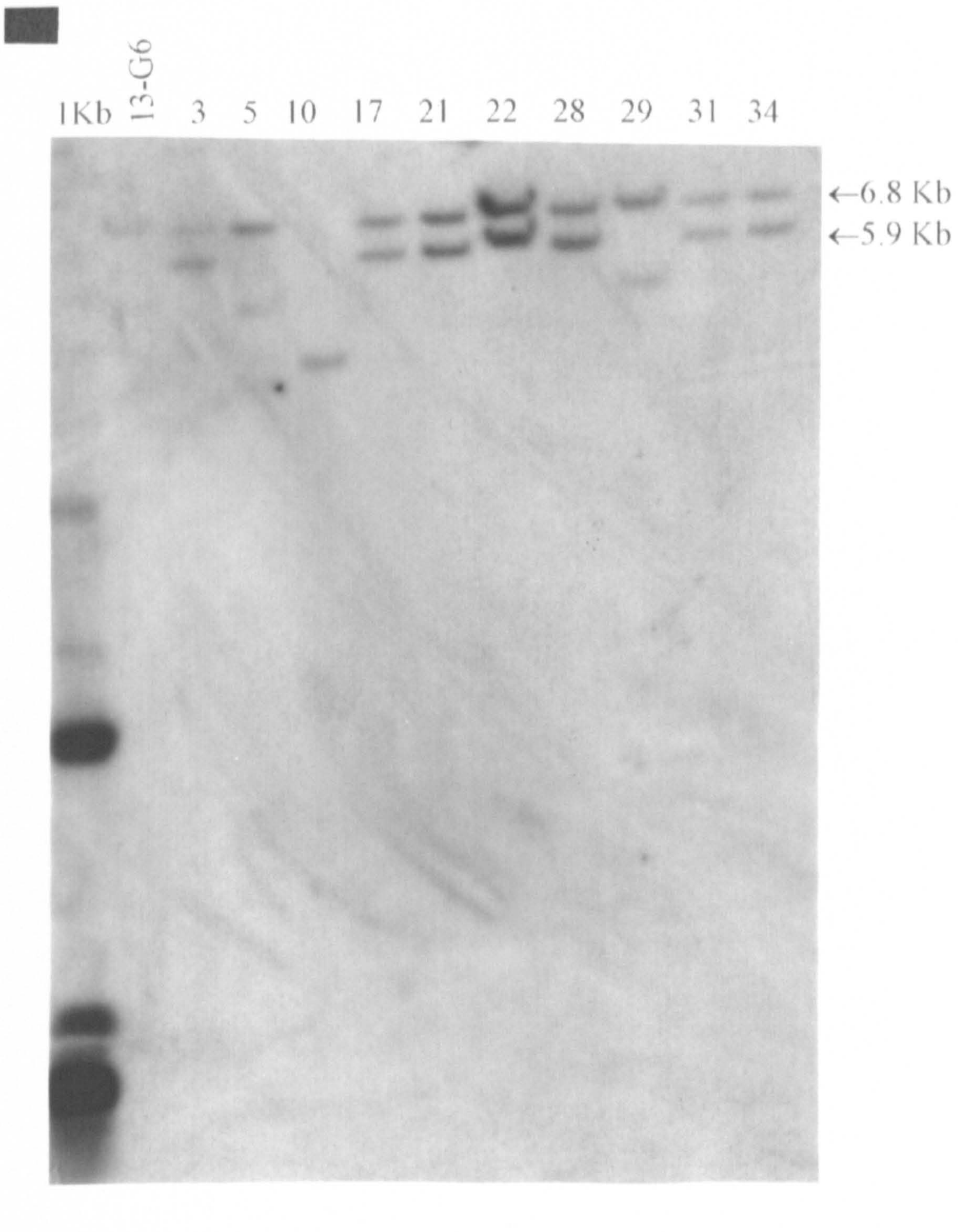


Figure 5.17a- Linearised map of the chromosomal target before Cre-mediated integration (parental)and after targeting with pHM3loxP. P corresponds to β-actin promoter.

Figure 5.17b- Detection of single-copy integrants corresponding to recombinant Cre protein mediated targeted integration of pHM3-loxP into the parental cell line, 13-G6. Genomic DNA (10 μg) from 10 independent transformants (observed as single copy integrants; see Figure 5.16) was each digested with *Nru*I and *Cla*I. Digested DNA was separated by electrophoresis on an agarose gel, and DNA fragments transferred onto a nylon membrane and probed with a radioactive human β-actin promoter fragment. The 6.8 Kb size indicated to the right corresponds to the parental fragment, whilst the 5.9 Kb fragment corresponds to the Cre-mediated targeted event (see Figure 5.17a). Lanes are labelled according to the number of the targeted clone. 1 Kb corresponds to molecular weight DNA marker.


1Kb
13-G6
3
5
10
17
21
22
28
29
31
34
←6.8 Kb
←5.9 Kb

recombination occurred after mitosis in only one descendant of a transfected cell, or if recombined and unrecombined cells mixed at replating or during subsequent growth. These unrecombined cells could have survived the G418 selection as a result of their close proximity to the targeted clones.

The results from the Southern hybridisation shows that the recombinant Cre protein can mediate the site-specific integration of pHM3-loxP to the chromosomal loxP site in cell line 13-G6. To characterise this targeted event further and also to show that the Cre protein mediated recombination event leads to the activation of the *lacZ* gene from the β-actin promoter and the subsequent loss or reduction of SAP expression, β-gal and SAP assays were determined.

5.9.2.3 Assay for β-gal and SAP expression from targeted clones

Single copy transgenes targeted to the same chromosomal loxP sites should express the same amount of recombinant proteins. Therefore, targeted clones were assayed for β-gal and SAP expression to see if they were reproducible and compared to the parental cell line (Result is shown in Table 5.5).

Results showed that the targeted event led to activation of *lacZ* (β-galactosidase activity) in all the clones analysed. With the exception of one transformant (clone 28 which showed a high level of β-galactosidase) all the clones analysed showed similar β -galactosidase activity. The high expression level of β-galactosidase in clone 28 was surprising however, because the integrated plasmid and the chromosomal content are identical in all cell lines. Although this clone was not examined in more detail (for instance to check if it showed the normal diploid karyotype) chromosomal duplication and/or tetraploidisation may account for this exceptional expression level. Only background β-galactosidase expression was observed in the parental cell line which did not contain the *lacZ* gene. The average level of SAP expression observed in the targeted clones was lower (although not significantly) than the level of SAP from the parental cell line. This high background of SAP expression in the targeted clones could be due to the mixture of unrecombined SAP expressing cells in the targeted clones as shown by Southern hybridisation (see section 5.8.2).

The result shows that the direct introduction of Cre protein into cells can mediate site-specific integration of an exogenous DNA to a chromosomal loxP site. In addition, the observation that lipofected recombinant Cre protein can catalyse site-specific

Transformant	β-galactosidase (mU)	SAP (arbitrary unit/10^6 cells/day)
13-G6 (parental)	-	2.45
3	6.2	1.9
5	4.9	1.45
10	1.7	2.51
17	8.2	1.21
21	7.9	1.38
22	4.7	2.75
28	16.1	1.58
29	5.8	3.85
31	9.6	2.1
34	7.4	1.78
Mean	**7.25**	**2.05**

Table 5.5- Estimation of β-galactosidase and SAP (SPR) activity from independent Cre-mediated targeted clones.

Transformants (1×10^6) cells were assayed for β-galactosidase activity (described in section 2.4.10) and SAP activity (described in section 2.4.8). Results of β-galactosidase and SAP activity are the average of assays from duplicate extracts from each transformant. Background β-galactosidase activity in 13-G6 parent cell line was subtracted from activities of the transformants. One unit of β-galactosidase hydrolyses 1 micromole of o-nitrophenyl-β-D-galactopyranoside (ONPG) to o-nitrophenol and galactose per minute at pH 7.5 at 37°C. Arbitrary unit of SAP is defined as the change in absorbance at A_{405nm} in 30 minutes (readings were done in duplicates).

recombination in cultured cells has important practical implications. Because of the ease of the procedure, it can be used for transformation of a target cell line with a battery of targeting plasmids to allow rapid generation of a series of isogenic cell lines. Moreover, removal of a marker gene flanked by loxP sites is simply accomplished by direct introduction of Cre protein, obviating the possibility of unwanted adventitious integration of a *cre* expression plasmid.

5.10 Summary

In this chapter, the Cre-loxP site-specific recombination system has been used to investigate the targeting of a promoterless *lacZ*/neo^r fusion gene to a chromosomal loxP site. A source of Cre recombinase to mediate the recombination event was provided by co-transfecting the targeting vector with either a *cre* expression plasmid or recombinant Cre protein purified from a baculovirus/insect system.

In order to use the Cre-loxP system for gene targeting in mammalian cell lines, it was necessary to establish an initial cell line which contains a single loxP site as a chromosomal target. Clearly, if multiple copies of the loxP sites are present in the chromosome, the Cre mediated recombination could lead to chromosomal translocation and re-arrangements which would be difficult to interpret. Southern blot analysis of transformants obtained using the Transfectam reagent, showed that this method of transfection favoured introduction of multiple copies of the loxP plasmid into the CHO chromosome. In contrast electroporation favoured the uptake of fewer copies of the foreign DNA. Furthermore, using 1μg of linearised DNA was found to favour the integration of one copy of the loxP plasmid into the CHO chromosome as shown by Southern blot analysis. Three cell lines containing a single copy of the loxP inserts were isolated by electroporating 1μg of linearised DNA into CHO cells.

The three cell lines (each containing a single loxP copy insert) showed different reporter SAP activity. Because the expression of the SAP gene in the plasmid pSAP-RDN/loxP is under the transcriptional control of β-actin promoter and integration of the SAP plasmid a random process, differences in SAP expression were a direct result of the effect of position of integration in the chromosome. Thus if integration disrupts the β-actin promoter, no detectable SAP expression would be observed, whilst a high level of SAP expression would indicate that integration had occurred in a favourable region in the chromosome for gene expression without disrupting the SAP gene.

Therefore, the difference in SAP activity between the 3 cell lines 13-G6 (highest), 1-G12 (medium) and 5-E10 (lowest) suggests that the loxP site had integrated in a favourable region for expression in 13-G6.

If the difference in SAP activity is the result of the chromosomal position effects, then the expression level of subsequent targeted genes to the same chromosomal position should correlate with SAP. To investigate this, the targeted vector, pHM3-loxP was co-transfected with a source of Cre recombinase into cell line 1-G12 and 13-G6. An estimate of the level of β-galactosidase activity from single copy integrants from the two different cell lines showed on average, approximately three fold higher β-galactosidase expression as a result of Cre mediated targeting of pHM3-loxP vector to 13-G6 than 1-G12. Because the targeting vector was directed to the same β-actin promoter which was initially driving the transcription of the SAP gene in both cell lines this result suggests that the environment of the chromosomal loxP site may have a direct effect on subsequent expression of the targeted gene. Thus in cell line 13-G6 where the loxP site integrated in a relatively favourable region in the CHO chromosome for expression as evidenced by the higher SAP activity, the same chromosomal position effect was observed on the targeted vector. This observation shows that Cre mediated targeted integration ensures that the same position effect (depending on the genomic location of the loxP target) is imposed on the incoming targeted gene. This is an important advantage in the use of the Cre-loxP site-specific recombination system as a tool for recombinant protein expression in mammalian cells. Thus if a loxP site is placed in a transcriptionally active region in the chromosome such as a region juxtaposed to a strong promoter or an enhancer region, Cre-mediated site-specific integration should favour the reproducible high level expression of the targeted gene. The SAP gene therefore provides a potential tool as an indicator for monitoring large numbers of transfected cells each containing a single loxP site in order to isolate cell line(s) which had integrated the loxP site in chromosomal regions favourable for the reproducible high level expression of subsequent targeted transgenes.

CHAPTER 6

Discussion

As described previously, large scale production of recombinant proteins such as recombinant antibodies have many applications in research and in the diagnosis and treatment of human diseases, and their efficient production will be of immense commercial importance. However, it is generally found that expression levels of recombinant proteins from stable clones isolated from a single transfection in mammalian cells show extreme variability, from a few nanograms of proteins to several micrograms per 10^6 cells/ml/day (Kaufman *et al.*, 1985). This is due to a highly variable number of copies integrated into the genome and from position effects on gene expression due to random integration hence extensive number of isolated clones must be screened in order to isolate a suitable donor for the high level expression of the recombinant protein. Thus, a system of introducing foreign DNA into mammalian cells that would eliminate variability in expression and which will allow high level expression of recombinant proteins would be beneficial for the industrial, large scale production of recombinant proteins.

The Cre-loxP site-specific recombination system (and other conservative site-specific recombination systems which have been shown to be active in heterologous eukaryotic cells; Golic and Linquist, 1989; O'Gorman *et al.*, 1991; Onouchi *et al.*, 1992) provides a strategy for the site-specific targeting of DNA into a single pre-selected region in the genome which should facilitate the construction of isogenic cell lines for the reproducible expression levels of recombinant proteins (Fukushige and Sauer, 1992). Such a strategy would be advantageous if the site to which the foreign gene is to be targeted is favourable for gene expression.

The objective of this work was to construct a mammalian cell line incorporating the Cre-loxP site-specific recombination system as a tool that would allow the site-specific targeting of exogenous DNA to pre-selected regions in the chromosomes which have been identified as favourable for gene expression. To achieve this, a chromosomal loxP target was constructed by linking a loxP site to a reporter gene encoding secreted alkaline phosphatase (SAP) to monitor gene expression in transfected cells. Transfected cells which contained a single copy of the loxP inserts were screened (using SAP expression) in order to isolate cell lines which had integrated the loxP site in regions of the chromosomes favourable for expression. Thus, subsequent targeting

of exogenous DNA to these sites, should allow the reproducible expression of high levels of recombinant proteins.

In this work, the Cre-loxP site-specific recombination reaction was initially studied *in vitro* using recombinant Cre protein which was purified from baculovirus-insect expression system with the aim of gaining a greater understanding into the mechanism of the reaction and also to identify irreversible substrates that would improve the efficiency of targeted integration in mammalian cells.

The Cre recombinase protein has been previously purified from *E. coli* (Abremski and Hoess, 1984). However, high level production of recombinant proteins have been reported in baculovirus-insect system by utilising the strong polyhedrin promoter to drive gene expression. For instance when placed under the transcriptional control of the polyhedrin promoter, human fibroblast interferon (Smith *et al.*, 1983), *E. coli* β-galactosidase (Pennock *et al.*, 1984) and human C-myc protein (Miyamoto *et al.*, 1985) were all expressed to very high levels, processed, modified and secreted appropriately in insect cells compared to the products synthesised in *E. coli* and *Saccharomyces cerevisiae* cells. For these reasons the Cre protein was expressed in the baculovirus/insect expression system.

Although the baculovirus/insect system allows the high level expression of recombinant proteins, purification from host endogenous proteins can be a time consuming task and a rate-limiting step in obtaining large quantities of material. To overcome this problem the Cre protein was expressed as a fusion with a hexa-histidine tag. Because histidine residues interact with a strong binding affinity for metal chelate adsorbents, a rapid single-step purification of the Cre protein from immobilised Ni^{++}-NTA affinity chromatography column was possible. Insect cells infected with a recombinant baculovirus (containing the hexa-histidine-*cre* fusion) expressed Cre protein, which was purified in a single-step from immobilised Ni^{++}-NTA affinity chromatography column. The purified protein migrated as a single polypeptide band on an SDS-polyacrylamide gel, with a molecular weight of approximately 38 kDa. Approximately 0.9 mg of Cre protein was produced from 1×10^9 insect infected cells compared to the levels of purified Cre protein from *E. coli*, reported to be 0.4 mg/g (wet weight) of cell culture (equivalent to approximately 2×10^{11} cells; Abremski and Hoess, 1985). It is possible that the high levels of Cre protein achieved with the baculovirus/insect system is due to the single purification step introduced, whereas a substantial amount of Cre protein was lost during the various purification steps with the *E. coli* system. The amount obtained from the baculovirus-insect system compares favourably with that of

c-myc protein produced from the same system (i.e. 1 μg of c-myc protein produced from 10^6 infected insect cells; Miyamoto *et al.*, 1985).

The Cre-loxP site-specific recombination reaction *in vitro* was verified by (i) Band shift assays and (ii) Cre-mediated intermolecular and intramolecular site-specific recombination at loxP sites:

(i) In Band shift (Revzins) assays, the Cre protein formed specific complexes with oligonucleotide sequences corresponding to loxP sites. The stoichiometry of the Cre-loxP interaction was consistent with two molecules of Cre bound to a full loxP site or one molecule bound to a half loxP site as reported by Mack *et al.*, (1992). The Cre-loxP interaction appeared to be temperature independent. However, it seems unlikely that an enzymatically catalysed reaction of this complexity would be entirely temperature independent. Furthermore, Drew *et al.*, (1985) proposed that the TpA sequence in the spacer region of the loxP site maybe predisposed to melting. Thus, it seems that the temperature effect maybe essential for other steps in the recombination process such as unwinding of the loxP site either during the pairing or strand exchange.

The initial binding of the recombinant Cre protein to its target loxP sites occurred very fast (under 30 seconds) making it difficult to determine the association/dissociation kinetics of the interaction. An "on-rate/off-rate" competition assay showed that a loxP site filled with two molecules of Cre was more stable than a half-filled loxP site, implying a possible protein-protein interaction between the Cre molecules that stabilises the complex formed at the loxP site probably through co-operative binding. The stability of the Cre-loxP binding interaction has important implications for the *in vivo* site-specific recombination in eukaryotic cells. This is because eukaryotic cells contain many DNA binding proteins which might be expected to compete with Cre for loxP sites so perhaps preventing productive encounters between Cre and its loxP sites. Thus the observation that the Cre-loxP complex is highly stable suggests that, not only can the complex be stable from non-specific interactions in the cell, but it can also be stable until it encounters another such complex for site-specific recombination. Although the rate of dissociation of the Cre molecules from the loxP site was not determined, the *in vitro* site-specific recombination can occur in less than a minute suggesting that the "off-rate" can be as rapid as the "on-rate".

One approach to investigate the dissociation/association constants of the Cre-loxP interaction would be the use of an automated analytical system; Biospecific Interaction Analysis (BIA) using BIAcore^{TM} (Pharmacia Biosensor). This technique uses

biosensors to detect and measure the kinetic interaction of any analyte-ligand reaction in real time (milliseconds). The principle depends on immobilising one component of the reaction, preferably the ligand on a sensor chip surface. As the analyte is passed over the chip and binds to the immobilised ligand, there is a change in mass on the sensor chip surface which correlates to a change in refractive index. The change in refractive index is measured as surface plasmon resonance (S.P.R.). The S.P.R. is monitored continuously so that chemical interactions between biochemicals can be studied in real time. For example, the BIAcore has provided useful information on the kinetics of the interaction of the Campath-1H series of antibodies with the CDW52 antigen; CHO-derived Campath-1H had an affinity constant (K_{aff}) 1.33-2.26 $\times$ 10^8 M- for CDW52, whereas Campath-1H aggregate material has a significantly different kinetic profile (slower on-rate and slower off-rate) from Campath-1H (Jon Ellis; pers. comm., GWR&D).

(ii) The recombinant Cre protein catalysed site-specific recombination *in vitro* without requiring any high energy cofactor(s) e.g., ATP or host protein(s) typically required for general recombination: recombination between two loxP sites in direct repeats relative to each other on a circular molecule gave rise to products each containing a single loxP site. These fragments then became substrates themselves because each of them contains an active loxP site (however the integration process is less favourable than the intramolecular excision event [Abremski *et al.*, 1983]). Recombination between two loxP sites which were inverted relative to each other resulted in the inversion of the intervening sequence flanked by the loxP sites. In the context of the substrates described, recombination was absolutely dependent on the presence of the recombinant Cre protein and functional loxP sites. The efficiency of recombination was dependent upon the concentration of the enzyme and its substrates and the duration of the reaction. Varying salt concentration did not significantly affect the efficiency of the reaction, however the effect of omitting Mg^{2+} was more significant than other salts on the efficiency of the reaction. Furthermore, cofactors such as BSA or spermidine were found to be essential for efficient recombination. The role of Mg^{2+} maybe to improve the stability of the Cre-loxP binding interaction because Cre protein has been observed to dimerise in the presence of Mg^{2+} (Abremski and Hoess, 1984), BSA may act to prevent denaturing of the protein as a result of non-specific interactions whilst spermidine may act to condense the DNA substrate bringing the loxP sites into close proximity for recombination to take place. These observations suggest that although the complexity and the large size of the chromatin compared to the loxP site may reduce the number of productive encounters between the loxP sites in the cell, the high concentrations of salts (especially polyamines) may act to condense the chromatin

structure and thereby bring the loxP sites into closer proximity for site-specific recombination to take place.

The Cre-loxP reaction is reversible, and excision is kinetically favoured over bi-molecular integration. Therefore, integrated products are inherently unstable in the presence of the recombinase and this can be a major drawback for the efficient use of the Cre-loxP site-specific recombination as a tool for targeted integration of transgenes in mammalian cells. The loxP site was manipulated *in vitro* in an attempt to identify substrates that could be used to improve the stability of the targeted integrations in mammalian cells. Substrate plasmids containing a wild-type loxP and a hybrid loxP (comprising half sites of loxP and frt) were used to investigate whether Cre would mediate intramolecular excision between the two sites. Although band shift assays showed that the Cre protein did not bind frt site, it is possible that Cre might have a weak interaction with the frt site which was not detected by the band shift conditions used. A hybrid loxP site would have reduced binding affinity for the recombinase, hence recombination event between a wild-type and a hybrid loxP site would be less probable; in effect excision would be less favourable because of weak association. Unfortunately, no recombination event was observed *in vitro* using the wild-type and the hybrid loxP site indicating that such hybrid loxP site could not be used as substrates for targeted integration. Rather than changing whole half sites, Alberts *et al.*, 1995; reported the introduction of selective base changes into the left (LE) and right (RE) 13 bp inverted repeats flanking the 8 bp spacer of the loxP site, such that recombination between a loxP site with a mutation in the LE and a site with a mutation in RE would produce two recombination product sites: one with mutation in neither the LE or RE (wild-type) site and the other with mutations in both LE and RE. If the LE and RE mutant sites has reduced binding affinity for the recombinase, subsequent recombination between the wild-type and mutant sites would be less probable hence reducing the frequency at which successfully integrated plasmids are subsequently excised. Recently, Araki *et al.*, (1997) reported site-directed DNA integration in ES cells using a pair of mutant loxP sites described by Alberts *et al.*, (1995) described above. Araki *et al.*, reported a 32-fold increase in site-specific integration using the mutant loxP sites compared to wild-type loxP sites. Such mutant loxP sites might be useful for the efficient Cre-mediated site-specific integration of transgenes in mammalian cells.

In conclusion, the results presented here shows that the recombinant Cre protein can mediate site-specific recombination *in vitro* with efficiency comparable to that of native Cre protein (Abremski *et al.*, 1983). Thus, although the recombinant Cre protein

contains a nuclear localisation signal and a hexa-histidine tag, they do not interfere with the recombinant Cre protein's activity. The histidine tag therefore acts as an independent domain which can be used for the detection of histidine-tagged Cre protein in Western blot assays or for the *in vivo* localisation of the Cre protein without the need to raise specific sera against the Cre protein. The nuclear localisation signal should increase the rate of nuclear import of the Cre protein. This is important for site-specific recombination events in mammalian cells where high concentration of the recombinase in the nucleus may be required for efficient recombination. The case with which the protein was expressed and purified suggests that this system can be used for the large scale production of Cre recombinase for site-specific recombination *in vitro* and *in vivo* and for further elucidation of the mechanism of site-specific recombination.

In the second part of this work, the Cre-loxP site-specfic recombination system was used as a tool that would allow the rapid targeting of any exogenous DNA to pre-selected regions in Chinese Hamster Ovary (CHO) genome for the reproducible expression of recombinant proteins. The Cre-mediated site-specific recombination in CHO cells requires a chromosomal loxP target for the integration event. Therefore, plasmid pSAP-RDN/loxP (which contains a loxP site between a β-actin promoter and SAP reporter gene) was used as the chromosomal loxP target such that stably transfected cells could be assayed for SAP expression in order to select cell lines which had integrated a single copy of the loxP plasmid in regions of the chromosomes favourable for gene expression. The loxP site was cloned in two orientations as determined by the directionality of the spacer region. In the one orientation (loxR), the loxP sequence was observed to have two ATG translational initiation codons upstream of the translational initiation codon of the SAP gene. The first upstream ATG codon was out of frame with the authentic translational initiation codon of the SAP gene and there is a purine (A) base at position -3, suggesting that this ATG codon was in a good context for translational initiation (Kozak, 1987). However no such ATG translational initiation codons were observed when the loxP site was cloned in the opposite orientation (loxL).

Plasmid pSAP-RDN/loxP (with the loxP site in both orientations) were independently transfected into CHO cells to determine whether a loxP site cloned in either orientation could have an effect on the expression of the SAP reporter gene. Results from the transfections showed that transfected cells containing loxR showed lower SAP expression compared to the transfected cells containing loxL. This observation showed that loxR cannot be used as the chromosomal loxP indicator because the upstream ATG codons in the loxP sequence interfered with SAP expression. For this reason, all

investigations were done using loxL which did not contain any other ATG codons apart from the authentic ATG codon of the SAP reporter gene.

For the Cre-mediated site-specific integration event, it was necessary to construct a cell line containing a single loxP site. This is because if multiple copies of the loxP sites are present in the chromosome, the Cre-mediated recombination event could lead to complex chromosomal translocations and re-arrangements. When the loxP plasmid (loxL), was transfected into CHO cells by the transfectam reagent, Southern blot analysis showed that transfected cells had incorporated multiple copies of the loxP plasmid into the genome showing that the transfectam reagent favoured the integration of many copies of the plasmid DNA into the genome. In contrast, electropration favoured the integration of fewer copies of the plasmid DNA into the CHO genome. Linearisation and low concentration of DNA favoured single copy integration. This observation is consistent with previous reports which showed that electroporation favours integration of one or few copies of exogenous DNA (Fromm *et al.*, 1986; Reiss *et al.*, 1986).

The decision to choose SAP expressing cell lines which have a single copy of the loxP insert resulted in only three cell lines out of forty which had a single loxP site. The low frequency of SAP expressing cell lines containing a single loxP site could be due to screening for cell lines which express medium to high levels of SAP activity. In addition, transfected cells contains a *dhfr* selectable marker under the transcriptional control of an attenuated SV40 early promoter, such that expression of DHFR should be less efficient than the reporter SAP. This bias for selection of transfected cells where the plasmid had integrated into transcriptionally active site and/or in multiple copies. Taken together, transfected cells containing single copy insertions may be under represented.

The difference in the reporter SAP expression levels observed in the three cell lines suggests that the loxP site had integrated into a different chromosomal region in each cell line. Thus the cell line expressing the highest SAP activity suggests that the loxP plasmid had integrated into a region in the chromosome favourable for gene expression whereas the cell line expressing the lowest SAP expression suggests that the loxP plasmid had integrated into a less favourable region. The SAP reporter gene could therefore be used to monitor the chromosomal environment of cell lines containing a single loxP site to establish cell lines where the loxP site is integrated in regions favourable for heterologous gene expression.

To investigate site-specific targeting of an exogenous DNA to the chromosomal loxP site, a promoterless *lac*Z/neo^r fusion plasmid containing a loxP site was co-transfected independently with both a *cre* expression plasmid and recombinant Cre protein (purified from baculovirus-insect system) to provide a source of the Cre recombinase. The targeting event was designed such that proper targeting of the promoterless *lac*Z/neo^r gene to the chromosomal loxP site (located between a β-actin promoter and the SAP gene) would displace the SAP gene from its' upstream β-actin promoter. Thus proper targeting is expected to give rise to individual G418 resistant transformants expressing similar levels of β-galactosidase and give loss or reduction of SAP activity. Random integration events on the other hand should give rise G418 resistant clones expressing different levels of β-galactosidase and the same level of SAP activity as the parental cell line.

The *cre* expression plasmid (which has been shown to encode active Cre recombinase in transient transfection assays) was used to mediate site-specific integration of the targeting vector (promoterless *lac*Z/neo^r fusion gene) to a chromosomal loxP site in cell line 1-G12. The advantage of using the promoterless *lac*Z/neo^r fusion gene to monitor Cre-mediated recombination event is that a large number of colonies can be rapidly screened in situ for β-gal expression. Results of β-gal staining of G418 resistant colonies isolated from this transfection showed that 26% of colonies resulting from Cre-mediated targeting event were positive for β-gal staining whilst only 4% of the colonies resulting from random integration activated β-galactosidase expression. The high frequency of β-galactosidase expression from the Cre-mediated recombination event indicates that the Cre-mediated reaction can direct the targeting of the promoterless *lac*Z/neo^r fusion gene to the chromosomal β-actin promoter site. Spontaneous activation of the *lac*Z gene as a result of illegitimate recombination to active region(s) in the chromosome such as near a promoter site occurred at low frequency.

A PCR analysis to investigate if the Cre-mediated event juxtaposed the targeting vector to a chromosomal loxP gave a lot of non-specific PCR products therefore a Southern hybridisation was used to investigate the recombination event. Results showed that eight out of twelve clones analysed from the Cre-mediated reaction showed one signal to suggest single-copy integrations. However, this signal was 1.2 kb smaller than would be expected for the targeted event suggesting a partial deletion of the targeting vector. Another Southern hybridisation using various restriction endonuclease digestions of genomic DNA from clones (isolated from the Cre-mediated reaction) to investigate the structural features all gave signals which migrated at

smaller sizes than expected. This result was a further indication that perhaps the targeting event had occurred but there was a deletion of a segment of the targeting vector probably by host cells' endo- and exo-nucleases activity prior to site-specific integration. Partial deletion of a targeting vector sequence as a result of Cre-mediated site-specific integration event has also been reported by Fukushige and Sauer, (1990); and Sauer and Henderson, (1989) although at much lower frequency in targeted clones than observed here. Although clones isolated from the Cre-mediated reaction were not analysed further, the isolation of clones which were: G418 resistant, positive for β-galactosidase expression (discuss later) and showed a reduction in SAP activity (discuss later) and also the result from Southern hybridisation analysis all suggests that the targeted event had occurred.

Because the Cre mediated reaction is freely reversible, and excision is kinetically favoured over bi-molecular integration, thus integrated products are inherently unstable in the presence of the recombinase. To avoid the possibility of the *cre* gene integrating into the CHO genome and being constitutively expressed, the targeted integration experiment was repeated using recombinant Cre protein which had been purified from baculovirus/insect expression system (described in chapter 3). The recombinant Cre protein mediated targeting event was investigated using cell line 13-G6. Results of the recombination event analysed by Southern hybridisation showed that ten out thirty-six clones isolated from the Cre-mediated recombination event showed one signal of the expected size suggesting that the recombinant Cre protein can mediate single-copy integration of the targeting vector in CHO cells. Other clones showed additional signals to suggest additional copies of the targeting vector randomly integrated elsewhere in the genome, presumably by illegitimate recombination.

Because site-specific integration using the *cre* expression plasmid and Cre protein were investigated in two different cell lines using two different methods of transfection in each case, the efficiency of the recombination cannot be compared. However, the direct introduction of Cre protein into cells shows a number of distinct advantages compared to using *cre* expression plasmids. First, use of purified protein rather than DNA eliminates the possibility of stable integration of the *cre* gene. For example, in the construction of targeted knockout mutations in embryonic stem cells it is often required to remove the selectable marker gene used to generate the knockout. This can be done by flanking the marker gene by directly repeated loxP sites in the targeting construct and later excising the marker gene in the resulting knockout cell line by transient expression of the *cre* gene (Sauer, 1993). Use of purified Cre protein eliminates the need to check marker gene "pop-out" colonies for the presence of the

cre gene. Second because direct protein transfer should be possible with many different transfer cell types, there is no need for individualised *cre* vectors having particular cell and tissue specific transcription and translation signals. Third, recombinase activity is temporally prescribed by the actual presence of functional Cre protein in the cell, thus averting problems for efficient integration of loxP targeting vectors arising from long term persistence (and expression) of DNA, as occurs in certain cells, such as striated muscle (Wolfe *et al.*, 1992). The baculovirus-insect expression system can therefore be used for the large scale production of recombinant Cre protein which can be delivered to many different cell types to investigate site-specific recombination.

As previously mentioned, the targeting event was designed such that proper targeting would give rise to G418 resistant transformants expressing β-galactosidase and give loss or reduction of SAP activity. None of the clones isolated from the Cre-mediated reaction in the two different cell lines showed a complete loss of SAP expression suggesting a residual SAP expression which could be due to transcriptional read-through from the β-actin promoter. With few exceptions, most of the clones showed a small but a reproducible level of β-galactosidase expression as a result of Cre-mediated targeted integration of the promoterless *lacZ*/neo^r fusion gene into the two independent single loxP cell lines. The level of β-galactosidase correlated with expression levels of the initial SAP reporter. Thus the cell line which showed the highest SAP expression also showed the highest β-galactosidase expression. This result indicates that the Cre-mediated targeting event ensures that the same position effect (depending on the genomic location of the loxP target) is imposed on the incoming targeting vector. This observation is consistent with previous report by Fukushige and Sauer, (1992). Fukushige and Sauer showed that Cre-mediated integration of exogenous DNA to different chromosomal positions resulted in expression levels of recombinant proteins characteristic of the particular chromosomal position. Thus the Cre-loxP system can be used for the controlled expression of recombinant protein by targeting the heterologous gene to a region in the chromosome which should favour gene expression.

Differences in gene expression observed in single copy integrants from one particular integration site must have come from influences other than copy number and position effect. One source of such variation is fluctuation in chromosome number from aneuploidy or tetrapoidy in cultured cells. Differences in methylation patterns which has been show to regulate gene expression (where methylated genes are inactive and the absence of methyl groups is associated with gene expression) may also affect gene expression (Doerfler, 1983).

In conclusion, the results presented here show that the Cre-loxP site-specific recombination system can be used as a tool for the targeting of any exogenous gene to a pre-selected region in the chromosome of CHO cells. Cre-mediated site-specific integration of a targeting vector to the chromosomal loxP site was investigated by co-transfecting a promoterless *lacZ*/neo^r fusion gene with a source of Cre recombinase into CHO cell lines containing loxP sites. The targeting event led to the activation of the promoterless *lacZ*/neo^r fusion gene from a β-actin promoter (upstream from the chromosomal loxP site). Both *cre* plasmid and recombinant Cre protein purified from baculovirus expression system successfully mediated the targeting event. The expression level of the targeted gene (β-galactosidase) correlated with the expression of SAP from the two independent loxP cell lines investigated, indicating that Cre-mediated targeting would ensure the same position effect (depending on the genomic location of the loxP target) on the incoming reporter gene. The ease and predictability of Cre-mediated single-copy transgene targeting suggests that this system would be beneficial for the large scale production of recombinant proteins because the Cre-loxP site-specific recombination system will provide a useful tool for the rapid generation of isogenic cell lines that should express reproducible levels of recombinant proteins. This is made possible because Cre can mediate the targeting of a single copy of transgenes to a pre-selected region in the chromosome which should eliminate extreme variability in expression levels (due to differing number of copies integrated into the genome and also from chromosomal "position effects") between individual clones isolated from a single transfection and therefore removing the need to screen large numbers of transformed cell lines for a suitable donor for recombinant protein expression.

Future Work

The work presented here shows that Cre-loxP site-specific recombination system can be used for both the *in vitro* and *in vivo* manipulation of transgenes. However, for the Cre-loxP system to be a useful tool for the reproducible high level expression of recombinant proteins in mammalian cells, the chromosomal region (containing a loxP site) to which the foreign gene is to be targeted must be favourable for gene expression. Thus a large number of cell lines each containing a single loxP site (linked to a reporter gene to monitor gene expression) must be screened in order to select cell lines where the loxP site had integrated into a favourable region for gene expression. Whilst the construction of CHO cell line containing a single loxP site for the targeting

event was successful, only three such cell lines were isolated when the loxP plasmid was transfected into CHO cells by electroporation. Although transfection by electroporation has been reported by Boggs *et al.*, (1986) to produce high frequency (79%) of single copy integrants, the low frequency of single loxP transformants (7.5%) obtained could be due to the different cell line used here than that reported by Boggs *et al.*,. It would be informative to investigate the electroporation conditions used further; such as comparing voltage versus transformation frequency and exposing high number of CHO cells (10^9/ml) to relatively low concentrations of linearised DNA to improve the frequency of single copy integration. Other electroporation parameters such as pulse length, size of DNA and conditions of the cell membrane could also be investigated for efficient transfection. Alternatively, large numbers of independent transfections could be done to obtain more transfected cell lines with the likelihood of obtaining more cell lines containing single loxP sites. However this would be a laborious way to achieve large numbers of loxP cell lines.

A more efficient way of constructing transformed cell lines containing single loxP sites could be the use a retrovirus vector mediated transfection. Retrovirus vectors (Palmer *et al.*, 1987; Miller and Rosman, 1989) can be genetically manipulated by replacing their genes with a cDNA containing the loxP sequence. Retrovirus vectors have the advantage of high infectivity (~ 100%) into a variety of cell types with the number of provirus being as low as one per cell. This may give rise to a high number of CHO transformants with a corresponding high numbers of CHO cell lines containing a single loxP site integrated randomly throughout the chromosome.

In order to improve the level of expression of a chromosomally integrated loxP plasmid, the loxP plasmid could be designed such that it is flanked by DNA elements known as MARs (Matrix attachment regions) or SARs (Scaffold attachment regions). These DNA elements act as matrices where chromatin loops of both interphase nuclei and metaphase chromosomes attach (Berenzney *et al.*, 1974; Mirkovitch *et al.*, 1984). However, these DNA elements are not only important for structural reasons as they have been shown to improve the expression levels of genes; Poljak *et al.*, (1994) and Phi-Van *et al.*, (1990). For instance, in the SARs/CAT expression system described by Poljak *et al.*, (1994), a SAR stimulatory effect of between 10-40 fold on CAT expression was observed when two minimal copies of Drosophila hsp70 or histone gene repeat SAR elements were used to flank an SV40 promoter-CAT reporter gene in a stable transfection. The SAR stimulatory effects however do not buffer the reporter gene from position effects, rather they consistently stimulate the average level of expression by over an order of magnitude and the effect appears not be tissue or gene-

specific. (There is contradicting evidence between the Poljak *et al.*, report and others [Stief *et al.*, 1989; Klehr *et al.*, 1991] that the minimal Drosophila SARs confer position-independent expression). A drawback to the stimulatory effect of SARs is that they appear to be blocked by GC-rich sequences when placed between SARs and the reporter gene.

Rather than random targeting of the loxP plasmid into mammalian cells to select for cell lines which have integrated the loxP plasmid in a favourable region for gene expression, a loxP plasmid could be targeted to a transcriptionally active regions such as an immunoglobulin enhancer region (by homologous recombination) so that subsequent targeted event would favour the high level expression of transgenes. However, this could only be applicable if the site-specific integration event is to be used for tissue-specific gene (e.g. Immunoglobulins) expression in the appropriate cell lines such as NS0.

An alternative approach to enhance gene expression would be to construct a loxP vector flanked by sequences comprising locus control region (LCR). This vector can then be targeted to the chromosomal LCR by homologous recombination to alleviate variation in gene expression due to random integration. LCR is a genetic element that contains dominant regulatory sequences that specify position-independent expression. In the LCR/MEL expression system described by Needham *et al.*, (1995), the human globin LCR was used to drive high-level, integration and position-independent expression of heterologous cDNAs in murine erythroleukaemia (MEL) cells. A vector was constructed to contain the human LCR upstream of the foreign cDNA expression cassette, which utilises the human β-globin promoter. This group expressed the human calcitonin receptor, a seven-transmembrane helix receptor, in this system achieving expression levels higher than previously achieved using an amplified CHO line and the baculovirus/insect cell system. Individual clones isolated following the use of LCR showed reproducible levels of gene expression. However, some studies have reported that the levels of expression obtained with the use of such LCR sequences are not higher than obtained following random integration and extensive screening (Grosveld *et al.*, 1987; Phi-Van *et al.*, 1990).

The Cre mediated site-specific recombination at loxP sites is reversible, and excision is kinetically favoured over bi-molecular integration, therefore Cre-mediated integration event is less efficient that the excision event. Furthermore, integrated products are inherently unstable in the presence of the recombinase. For these reasons, it is important to deliver the Cre recombinase efficiently and transiently for intermolecular

(integration) event to occur. Various approaches have been applied to improve the efficiency of gene targeting that abolishes post-integration events: The use of transient *cre* expression plasmid or Cre protein has been described in this work and previously by Sauer and Henderson, (1990), Gu *et al.*, (1993), Baubonis and Sauer (1993).

Another approach that could be used to abolish post-integration recombination would be insertional inactivation of *cre*. For instances, a promoterless loxP-neo^r (linked to the transgene) construct could be targeted to a transformed parental cell line containing a promoter-loxP-*cre* construct. The parental cell line would synthesis Cre, which would mediate the insertion of the targeting vector into the genomic loxP site to produce promoter-loxP-neo-loxP-cre linkage. Integration of loxP-neo^r would displace the promoter from the *cre* to terminate *cre* transcription and fuses the promoter to the neo^r gene to confer a neomycin resistance phenotype. This strategy has been tested in plants (Alberts *et al.*, 1995). However constitutive expression of Cre could accumulate a large amount of the recombinase which could be slow to deplete. Alternatively or in addition, promoter displacement may fail to abolish Cre production completely, but may leave a residual level of *de novo* synthesis. Residual Cre activity could delete inserts flanked by wild-type loxP sites. However, this would not be a problem if mutant loxP sites (discussed above) are used.

Appendices

Appendix 1: Nucleotide sequence coding for cre which is modified to include nuclear localisation sequence (NLS). The transcriptional initiation (ATG) and termination (TAG) codons are shown in bold and the NLS sequence is underlined.

TCGACC**ATG**C CCAAGAAGAA GAGGAAGGTG TCCAATTTAC TGACCGTACA
CCAAAATTTG CCTGCATTAC CGGTCGATGC AACGAGTGAT GAGGTTCGCA
AGAACCTGAT GGACATGTTC AGGGATCGCC AGGCGTTTTC TGAGCATACC
TGGAAAATGC TTCTGTCCGT TTGCCGGTCG TGGGCGGCAT GGTGCAAGTT
GAATAACCGG AAATGGTTTC CCGCAGAACC TGAAGATGTT CGCGATTATC
TTCTATATCT TCAGGCGCGC GGTCTGGCAG TAAAAACTAT CCAGCAACAT
TTGGGCCAGC TAAACATGCT TCATCGTCGG TCCGGGCTGC CACGACCAAG
TGACAGCAAT GCTGTTTCAC TGGTTATGCG GCGGATCCGA AAAGAAAACG
TTGATGCCGG TGAACGTGCA AAACAGGCTC TAGCGTTCGA ACGCACTGAT
TTCGACCAGG TTCGTTCACT CATGGAAAAT AGCGATCGCT GCCAGGATAT
ACGTAATCTG GCATTTCTGG GGATTGCTTA TAACACCCTG TTACGTATAG
CCGAAATTGC CAGGATCAGG GTTAAAGATA TCTCACGTAC TGACGGTGGG
AGAATGTTAA TCCATATTGG CAGAACGAAA ACGCTGGTTA GCACCGCAGG
TGTAGAGAAG GCACTTAGCC TGGGGGTAAC TAAACTGGTC GAGCGATGGA
TTTCCGTCTC TGGTGTAGCT GATGATCCGA ATAACTACCT GTTTTGCCGG
GTCAGAAAAA ATGGTGTTGC CGCGCCATCT GCCACCAGCC AGCTATCAAC
TCGCGCCCTG GAAGGGATTT TTGAAGCAAC TCATCGATTG ATTTACGGCG
CTAAGGATGA CTCTGGTCAG AGATACCTGG CCTGGTCTGG ACACAGTGCC
CGTGTCGGAG CCGCGCGAGA TATGGCCCGC GCTGGAGTTT CAATACCGGA
GATCATGCAA GCTGGTGGCT GGACCAATGT AAATATTGTC ATGAACTATA
TCCGTAACCT GGATAGTGAA ACAGGGGCAA TGGTGCGCCT GCTGGAAGAT
GGCGAT**TAG**C CATTAACGCG TAAATGATTG CTATAATTAG TTGATA

Appendix 2: Sequence of oligonucleotide primers (synthesised by H. Spence, WRL [now GWR&D]) used in either PCR or Sequencing reactions.

CB1741: GACAAGAGATCT**ATG**CCCAAGAAGAAGAGGAAGG

CB1742: GACAAGAGATCT**CTA**ATCGCCATCTTCCAGC

CB1741 anneals at the 5' end at position 7bp-28bp and CB1742 anneals at the 3' end at position 1059bp-1041bp of the *cre* coding sequence (appendix 1). CB1741 and CB1742 were used in a PCR reaction to amplify *cre* gene from pMC1-Cre. The ATG initiation and the CTA (anti-sense) termination codons in CB1741 and CB1742 respectively are in bold. The introduced *Bgl*II sites are underlined.

BK59: CGTGTTTTAATACGCCGGACC

BK59 anneals upstream of *Kpn*I site at position 4967bp-4946bp in pAcH6N1 (Figure 3.3). BK59 and CB1741 (above) were used in a PCR reaction to screen recombinant baculovirus.

BK62: CATACCGTCCCACCATCGGG

BK62 anneals upstream of *Bam*HI site at position 5062bp-5082bp in pAcH6N1 (Figure 3.3). BK62 and CB1742 (above) were used in a PCR reaction to screen recombinant baculovirus.

B142: GACGTCGAATTCTTTCCATTCGCCATTCAGGCT

B143: GACGTCAAGCTTGCGCAACGCAATTAATGTGA

B142 anneals at the 5' end at position 460bp-480bp, whilst B143 anneals at the 3' end at position 940bp-920bp in pBluescript II KS (-) (Stratagene). B142 and B143 were used in a PCR reaction to amplify *lac*Z gene from ΔpBS KS II (section 3.6). The introduced *Eco*RI site in B142 and *Hind*III site in B143 are underlined.

CB2076: CATTATTATCATGACATTAACCT

CB2076 anneals upstream of EcoRI site at position 4300bp-4323bp in pAT-lacZloxP2 (fig 3.7). CB2076 was used in a sequencing reaction with template DNA from product A (Figure 3.8).

CB1736: GCTATTCTCGCAGGATCAGTCGACC

CB1738: CCAGTTTGTCCTTCTTCTGCCCTTT

CB1736 anneals upstream of the *Hind*III site at position 10663bp-10688bp whilst CB1738 anneals downstream of the *Hind*III site at position 263bp-288bp in pSAP-RDN (Figure 5.3a). CB1736 and CB1738 were used in a PCR reaction to confirm the presence of intact loxP site in transfected cells.

CB1739: ATAATTCGCGTCTGGCCTTCCTGTA

CB1739 anneals upstream of the *Cla*I site at position 3169bp-3194bp in pHM3 (Figure 5.12). CB1739 and CB1736 (above) were used in a PCR reaction to confirm targeting event.

Appendix 3: Composition of Boehringer Mannheim restriction enzyme buffers A, B, L, M and H.

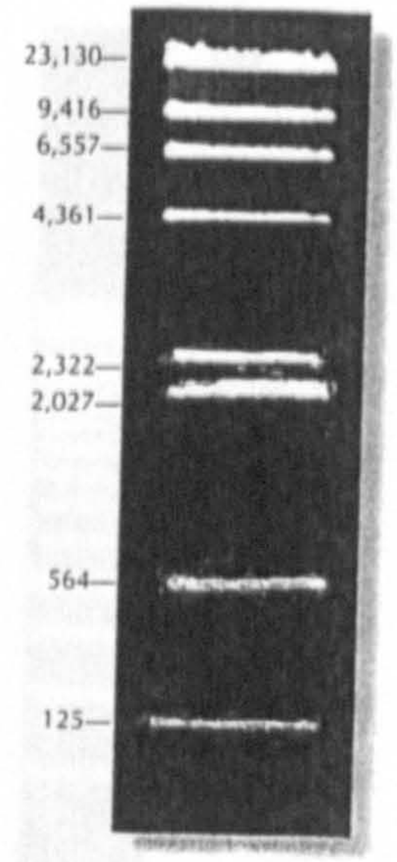
Buffer Composition	A	B	L	M	H
Tris acetate	33	-	-	-	-
Tris-HCl	-	10	10	10	50
Mg-acetate	10	-	-	-	-
$MgCl_2$	-	5	10	10	10
K-acetate	66	-	-	-	-
NaCl	-	100	-	50	100
Dithioerythritol	-	-	1	1	1
Dithiothreitol	0.5	-	-	-	-
2-Mercaptoethanol	-	1	-	-	-
pH at 37°C	7.9	8.0	7.5	7.5	7.5

Final concentration in nM (following 1:10 dilution of 10× buffer).

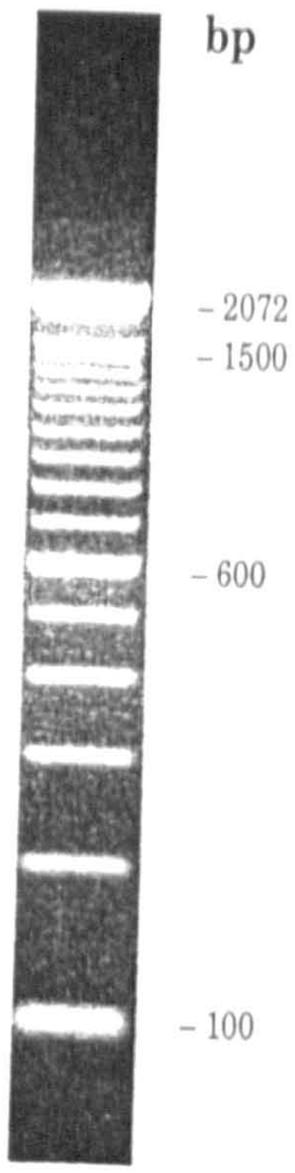
Appendix 4: Oligonucleotide sequences synthesised (by H. Spence; WRL [now GWR&D]) as sense and antisense strands to correspond to loxP; Frt; loxP/Frt hybrids.

LoxP	TAATATAACTTCGTATA	**ATGTATGC**	TATACGAAGTTATTAGG
	ATTATATTGAAGCATAT	**TACATACG**	ATATGCTTCAATAATCC
Frt	taatgaagttcctatac	**tttctaga**	gaataggaacttctagg
	attacttcaaggatatg	**aaagatct**	cttatccttgaagatcc
LoxP/Frt hybrid			
(loxP spacer)	TAATATAACTTCGTATA	**ATGTATGC**	gaataggaacttctagg
	ATTATATTGAAGCATAT	**TACATACG**	cttatccttgaagatcc
LoxP/Frt hybrid			
(frt spacer)	TAATATAACTTCGTATA	**tttctaga**	gaataggaacttctagg
	ATTATATTGAAGCATAT	**aaagatct**	cttatccttgaagatcc

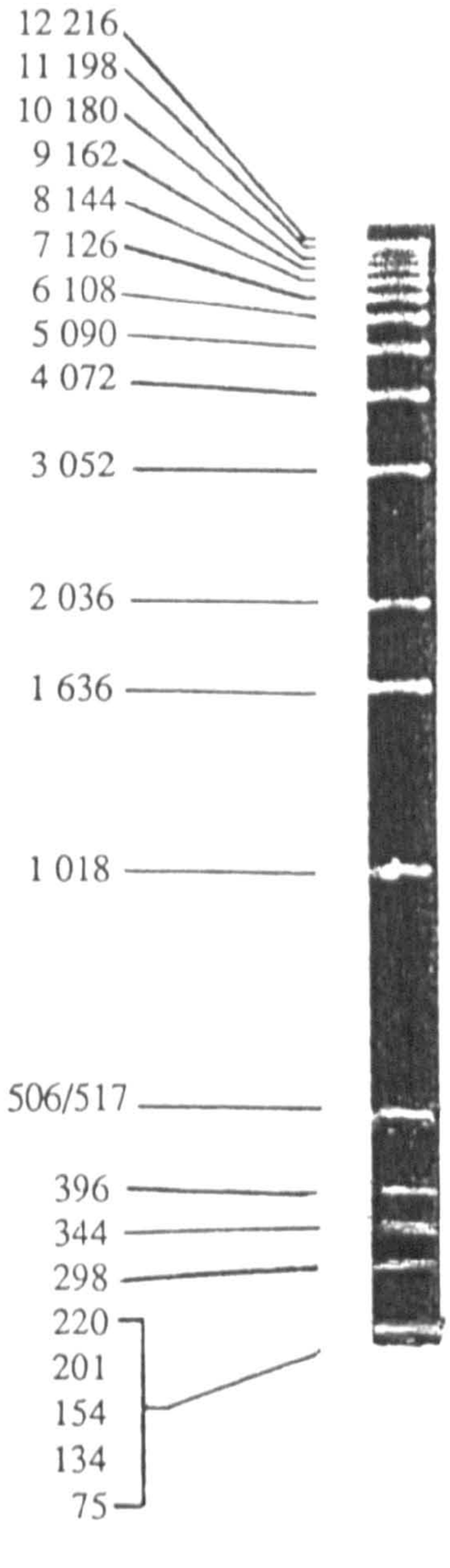
Appendix 5: Lambda (λ) DNA *Hin*d III digest (GIBCO BRL) molecular weight size markers for agarose gel electrophoresis. The sizes of fragments (in base pairs) and their expected pattern following agarose gel separation are indicated below.



Appendix 6: 100 bp ladder (GIBCO BRL) molecular weight size markers for agarose gel electrophoresis. The sizes of fragments (in base pairs) and their expected pattern following agarose gel separation are indicated below.

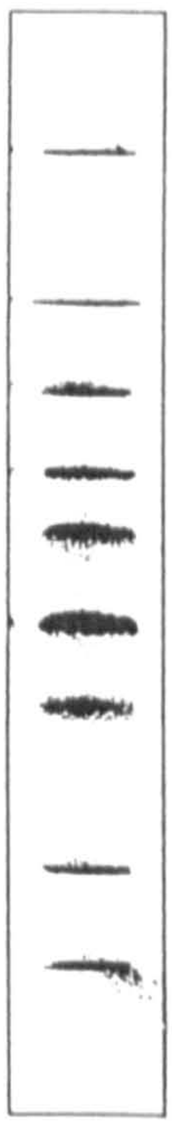


Appendix 7: 1 kb ladder DNA (GIBCO BRL) molecular weight size markers for agarose gel electrophoresis. The sizes of fragments (in base pairs) and their expected pattern following agarose gel separation are indicated below.



Appendix 8: Sea Blue (Novex) molecular weight size markers for PAGE.
The sizes of proteins (in kDa) and their expected pattern following separation on a 10-20% acrylamide gel are indicated below.

Protein	Approx. mol. weight (kDa)
myosin	250
BSA	98
glutamic dehydrogenase	64
alcohol dehydrogenase	50
carbonic anhydrase	36
myoglobin	30
lysozyme	16
aprotinin	6
insulin, B chain	4



Appendix 9: The degenerate codes for DNA nucleotide sequences

R= A or G

Y= C or T

M= A or C

K= G or T

S= G or C

W= A or T

H= A or T or C

B= G or T or C

G= G or A or T

N= A or G or C or T

V= G or A or C

References

Abuin, A. and Bradley, A. (1996). Recycling selectable markers in mouse Embryonic Stem Cells. Mol. Cell Biol. **16**; 1851-1856.

Abken, H. and Reifenrath, B. (1992). A procedure to standardise CAT reporter gene assay. Nucl. Acids Res. **20**; 3527.

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.

Abremski, K. and Hoess, R. (1984). Bacteriophage P1. Site-Specific Recombination. J. Biol. Chem. **259**; 1509-1514.

Abremski, K. and Hoess, R. H. (1992). Evidence for a 2nd. Conserved Arginine Residue In the Integrase Family of Recombination Proteins. Protein Engng. **5**; 87-91.

Adams, D. E., Bliska, J. B. and Cozzarelli, N. R. (1992). Cre-lox Recombination in *Escherichia coli* cells. Mechanistic Differences from the in vitro Reaction. J. Mol. Biol. **226**; 661-673.

Adang, M. J. and Miller, L. K. (1982). Molecular cloning of DNA complementary to messenger RNA of the baculovirus *Autographa californica* nuclear polyhedrosis virus: Location and gene products of RNA transcription found late in infection. J. Virol. **44**; 782-793.

Alam, J. and Cook, J. L. (1990). Reporter genes: application to the study of mammalian gene transcription. Anal. Biochem. **188**; 245-254.

Albert, H., Dale, E. C., Lee, E. and Ow, D. W. (1995). Site-specific integration of DNA into wild-type and mutant loxP sites placed in the plant genome. The Plant Journal **7**; 649-659.

Amrein, K. E., Takacs, B., Steiger, M., Molnos, J., Flint, N. A. and Burns, P. (1995). Purification and characterisation of recombinant human p50csk protein-tyrosine kinase from *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. Proc. Natl. Acad. Sci. USA. **92**; 1048-1052.

An, G., Hidaka, K. and Siminovitch, L. (1982). Expression of bacterial β-galactosidase in animal cells. Mol. Cell. Biol. **2**; 1628-1632.

Anderson, W. F., Ohlendorf, D. H., Tekeda, Y., Matthers, B. W. (1981). Structure of the Cre repressor from bacteriophage-lambda and it's interaction with DNA. Nature **290**; 754-758.

Andrews, B. J., Proteau, G. A., Beatty, L. G. and Sadowski, P. D. (1985). The Flp Recombinase of the 2-MU Circle DNA of yeast- Interaction with it's target sequence. Cell **40**; 795-803.

Araki, K., Araki, M. and Yamamura, K. (1997). Targeted integration of DNA using mutant lox sites in embryonic stem cells. Nucl. Acids Res. **25**; 868-872.

Argos, P., Landy, A., Abremski, K., Egan, J. B., Ljungquist, E. H., Hoess, R. H., Kahn, M. L., Kalionis, B., Narayana, S. V. L., Pierson, L. S., Sternberg, N. and Leong, J. M. (1986). The integrase family of site-specific integrases: regional similarities and global diversity. EMBO J. **5**; 433-440.

Aris, J. P., Basta, P. V., Holmes, W. D., Ballas, L. M., Moomaw, C., Rankl, N. B., Blobel, G., Loomis, C. R. and Burns, D. J. (1993). Molecular and biochemical characterization of a recombinant human PKC-delta family member. Biochim. Biophys. Acta. **1174**;171-181.

Austin, S., Ziese, M. and Sternberg, N. (1981). A novel role for site-specific recombination in maintenance of bacterial replicons. Cell **25**; 729-736.

Bächi, B. and Arber, W. (1977). Physical mapping of *Bgl*II, *Bam*HI, *Eco*RI, *Hind*III and *Pst*I restriction fragments of bacteriophage PI DNA. Mol. Gen. Genet. **153**; 311-324.

Baker, M. D., Murraldo, H. and Shulman, M. J. (1988). Expression of an immunoglobulin kappa light-chain gene in lymphoid cells using a Bovine Papillomavirus (BPV-1) vector. Gene **69**; 349-355.

Bancroft, I., Jones, J. D. G. and Dean, C. (1993). Heterologous transposon tagging of the DRL1 locus in *A. rabidopsin*. Plant Cell **5**; 631-638.

Baubonis, W. and Sauer, B. (1993). Genomic targeting with purified Cre recombinase. Nucl. Acids Res. **21**; 2025-2029.

Bayley, C. C., Morgan, M., Dale, E. C. and Ow, D. W. (1992). Exchange of gene activity in transgenic plants catalysed by the Cre-loxP site-specific recombination system. Plant Molecular Biology **18**; 353-361.

Beames, B., Braunagel, S., Summers, M. D., and Lanford, R. E. (1991). Polyhedron initiator codon altered to AUU yields unexpected fusion protein from a baculovirus vector. Biotechniques **11**; 378-383.

Bebbington, C. R., Renner, G., Thomson, S., King, D., Abrams, D. and Yarranton, G. T. (1992). High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. Bio/Technology. **10**; 169-175.

Bebbington, C. R. and Hentschel, C. C. G. (1985). The expression of recombinant DNA products in mammalian cells. Tibtech. **3**; 314-317.

Bebbington, C. R. and Hentschel, C. C. G. (1987). DNA Cloning. ed. Glower, D. M. IRL. Press, New York. Vol 3, pp163-171.

Belyaev, A. S. and Roy, P. (1993). Development of baculovirus triple and quadruple expression vectors: co-expression of three or four blue tongue virus proteins and the synthesis of blue tongue virus-like particles in insect cells. Nuc. Acids. Res. **21**; 1219-1223.

Berg, O. G., Winter, R. B. and Von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry **20**; 6929-6948.

Berenzney, R. and Coffey, D. (1974). Identification of a nuclear protein matrix. Biochem. Biophys. Res. Commun. **60**; 1410-1419.

Berger, J., Howard, A. D., Gerber, L., Cullen, B. R. and Udenfriend, S. (1987). Expression of active membrane-bound human placental alkaline phosphatase by transfected simian cells. Proc. Natl. Acad. Sci. USA. **84**; 4885-4889.

Bertani, G. (1951). Studies in lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J. Bacteriol. **62**; 293-300.

Bertini, L. E., and Six, E. W. (1988). The P2-like phages and their parasite, P4, in: The Bacteriophages (R. Calender, ed.), Plenum, New York. Vol. 2, pp 73-143.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7; 1513-1523.

Bishop, D. H. L. (1992). Baculovirus expression vectors. Seminars in Virol **3**; 253-264.

Boggs, S. S., Gregg, R. G., Borenstein, N. and Smithies, O. (1986). Efficient Transformation and Frequent Single-site Single-copy Insertion of DNA can be obtained in Mouse Erythroleukemia Cells Transformed by Electroporation. Exp. Hematol. **14**; 988-994.

Bonning, B. C., Roelvink, P. V., Vlak, J. M., Possee, R. D. and Hammock, B. D. (1994). Superior expression of juvenile hormone esterase and beta-galactosidase from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus compared to the p10 protein and polyhedrin promoters. J. Gen. Virol. **75**; 1551-1556.

Boocock, M. R., Brown, J. L. and Sherratt, D. J. (1985). Structural and catalytic properties of specific complexes between TN3 Resolvase and the recombination sites, Res. Biochem. Soc. Trans. **14**; 214-216.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of Protein-Dye binding. Anal. Biochem **72**; 248-254.

Broach, J. R. and Hicks, J. B. (1980). Replication and Recombination Functions Associated with the Yeast Plasmid 2μ Circle. Cell **21**; 501-508.

Brown, M. and Faulkner, P. (1977). A plaque assay for nuclear polyhedrosis viruses using a solid overlay. J. Gen. Virol **36**; 361-364.

Buchman, A. R. and Berg, P. (1988). Comparison of intron-independent Gene-Expression. Mol. Cell. Biol. **8**; 4395-4405

Capecchi, M. R. (1989). The new mouse genetics; altering the genome by gene targeting. Trends Genet **5**; 70-76.

Carbonell, L., Hodge, M. R., Tomalski, M. D. and Miller, L. K. (1988). Synthesis of a gene coding for an insect specific scorpion toxin and attempts to express it using baculovirus vectors. Gene **73**; 409-418.

Carroll, A. R., Rowlands, D. J. and Clarke, B. E. (1992). Synthesis and secretion of a functional antibody in a vaccinia virus expression system. Mol. Immunol. **29**; 821-827.

Chesney, R. H., Scott, J. R. and Vapnek, D. (1979). Integration of the plasmid prophages P1 and P7 into the chromosome of *Escherichia coli*. J. Mol Biol. **130**; 161-173.

Chu, G. and Berg, P. (1985). Rapid assay for detection of *Escherichia coli* xanthine-guanine phosphoribosyl transferase activity in transfected cells. Nucl. Acids Res. **13**; 2921-2930.

Chu, G., Hayakawa, H. and Berg, P. (1987). Electroporation for the efficient transfection of mammalian cells with DNA. Nucl. Acids Res. **15**; 1311-1326.

Chuck, G., Robbins, T., Nijjar, C., Ralston, E., Courtney-Gutterson, N. and Dooner, H. K. (1993). Tagging and cloning of a petunia flower colour gene with the maize transposable element Activator. Plant Cell **5**; 371-378.

Cockett, M. I., Bebbington, C. R. and Yarranton, G. T. (1990). High-level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. Bio/Technology **8**; 662-667.

Colbere-Gerapin, F. F., Chousterman, S., Horodniceanu, F., Kourilsky, P., Garapin, A-C. Cloning of the active thymidine kinase gene of herpes simplex virus type 1 in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. **76**; 3755

Cole, E. S., Lee, K., Lauziere, K., Kelton, C., Chappel, S., Weintrails, B., Ferrara, D., Peterson, P., Bernasconi, R., Edmunds, T., Richards, S., Dickrell, L., Kleeman, J. M., McPherson, J. M. and Pratt, B. M. (1993). Recombinant human thyroid stimulating hormone: development of a biotechnology product for detection of thyroid carcinoma. BioTechnology **11**; 1014-1024.

Copp, A. J. (1995). The tale that wags the Embryo-Reply. Trends in Genetics **11**; 298-299.

Craig, N. C. (1988). The mechanism of site-specific recombination. Annu. Rev. Genet. **22**; 77-105.

Craig, N. L. and Nash, H. A. (1983). The mechanism of Phage λ Site-Specific Recombination : Site-Specific Breakage of DNA by Int Topoisomerase. Cell **35**; 795-803.

Cress, D. E. and Kline, B. C. (1976). Isolation and characterisation of *Escherichia coli* chromosomal mutants affecting plasmid copy number. J. Bacteriol. **125**; 635-642.

Dale, E. C. and Ow, D. W. (1990). Intra- and intermolecular site-specific recombination in plants cells mediated by bacteriophage P1 recombinase. Gene. **91**; 79-85.

Dale, E. C. and Ow, D. W. (1991). Gene transfer with subsequent removal of the gene from the host genome. Proc. Natl. Acad. Sci. USA. **88**; 10558-10562.

Denhardt, D. T. (1966). A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**; 641-646.

Davies, A. H., Jowett, J. B. M. and Jones, I. M. (1993). Recombinant Baculovirus Vectors Expressing Glutathione-S-Transferase Fusion Proteins. BIO/TECHNOLOGY **11**; 933-936.

Davies, A. H. (1994). Current methods for manipulating baculoviruses. Bio/Technology. **12**; 47-50.

Depew, R. E. and Wang, J. C. (1975). Conformational fluctuations of DNA helix. Proc. Natl. Acad. Sci. USA **72**; 4275-4279.

Doerfler, W. (1983). DNA methylation and Gene Activity. Ann. Rev. Biochem. **52**; 93-124.

Doerfler, W. (1986). Expression of the *Autographa californica* nuclear polyhedrosis virus genome in insect cells: homologous viral and heterologous vertebrate genes in the baculovirus vector system. Curr. Top. Microbiol. Immunol. **131**; 51-68.

Dorai, H. and Moore, G. P. (1987). The effect of dihydrofolate reductase-mediated gene amplification on the expression of transfected immunoglobulin genes. J. Immunol. **139**; 4232-4241.

Drew, H. R., Weeks, J. R. and Travers, A. (1985). Negative supercoiling induces spontaneous unwinding of a bacterial promoter. EMBO J. **4**; 1025-1032.

Eglitis, M. A. (1991). Positive selectable markers for use with mammalian cells in culture. Hum. Gene Ther. **2**; 195-207.

Elledge, S. J., Mulligan, J. T., Ramer, S. W., Spotswood, M. and Davis, R. W. (1991). λ YES: a functional cDNA expression vector for the isolation of genes by complementation of yeast and *Escherichia coli* mutations. Proc. Natn. Acad. Sci. USA. **88**; 1731-1735.

Elroy-Stein, O., Fuerst, T. R. and Moss, B. (1989). Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. Proc. Natl. Acad. Sci. USA. **86**; 6126-6130.

Epstein, C. J. (1986). The Consequences of Chromosome Imbalance: Principles, Mechanism and Models (Cambridge Univ. Press).

Evans, H. F. (1986). Ecology and epizootiology of baculoviruses. In the Biology of Baculoviruses, Vol. 2, eds. Granados, R. R. and Federici, B A. Boca Raton, Fla: CRC Press. pp 89-132.

Eves, E. M. and Farber, R. A. (1983). Expression of recessive Aprt-mutations in mouse CAK cells resulting from chromosome loss and duplication. Somat. Cell. Genet. **9**; 771-778.

Feng, J. A., Johnson, R. C. and Dickerson, R. E. (1994). Hin Recombinase Bound to DNA: The Origin of Specificity in Major and Minor Groove Interactions. Science **263**; 348-355.

Foecking, M. K. and Hofstetter, H. (1986). Powerful and versatile enhancer-promoter unit for mammalian expression vectors. Gene **45**; 101-105.

Fouser, L. A., Swanberg, S. L., Lin, B. Y., Benedict, M., Kelleher, K., Cumming, D. A. and Reidel, G. E. (1992). High level expression of a chemeric anti-ganglioside GD2 antibody: genomic kappa sequences improve expression in COS and CHO cells. Bio/Technology **10**; 1121-1127.

Fraser, M. J. (1986). Ultrastructural observation of virion maturation in Autographa Californica nuclear polyhedrosis virus infected *Spodoptera frugiperda* cell culture. J. Ultrastruct. Mol. Struct. Res. **95**; 189-195.

Fried, M. and Crothers, D. M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nuc. Acids Res. **9**; 6505-6525.

Fromm, M. E., Taylor, L. P. and Walbot, V. (1986). Stable transformation of maize after gene transfer by electroporation. Nature **314**; 791-793.

Fry, M. (1983). Enzymes of Nucleic Acid Synthesis and Modification, S. T. Jacobs, Ed. (CRC Press, Boca Raton, FL) pp 39-92.

Fukushige, S. and Sauer, B. (1992). Genomic targeting with a positive-selection loxP integration vector allows highly reproducible gene expression in mammalian cells. Proc. Natl. Acad. Sci. USA. **89**; 7905-7909.

Gage, P. J., Sauer, B., Levine, M. and Glorioso, J. C. (1992). A cell-free recombination system for site-specific integration of multigenic shuttle plasmids into the herpes simplex virus type 1 genome. J. Virol. **66**; 5509-5515.

Garner, M. M. and Rerzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to componencts of the *Escherichia coli* lactose operon regulatory system. Nuc. Acids Res. **9**; 3047-3060.

Gasser, C. S. and Fraley, R. T. (1989). Genetically engineered plants for crop improvement. Science **244**; 1293-1299.

Gellert, M. and Nash, H. (1987). Communications between segments of DNA during site-specific recombination. Nature **325**; 401-404.

Godeau, F., Saucier, C. and Kourilsky, P. (1992). Replication inhibition by nucleoside analogues of a recombinant Autographa Californica multicapsid nuclear polyhedrosis virus harbouring the herpes thymidine kinase gene driven by the IE-1 (O) promoter: a new way to select recombinant baculoviruses. Nucl. Acids Res. **20**; 6239-6246.

Godwin, A. R., Bollag, R. T., Christie, D. and Liskay, R. M. (1994). Spontaneous and restriction enzyme-induced chromosomal recombination in mammalian cells. Proc. Natl. Acad. Sci. USA. **91**; 12554-12558.

Golic, K. G. and Lindquist, S. (1989). The FLP Recombinase of Yeast Catalyses Site-Specific Recombination in the Drosophila Genome. Cell **59**; 499-509.

Goodman, R. M., Hauptli, H., Crossway, K. A. and Knauf, V. C. (1987). Gene-transfer in crop improvement. Science **236**; 48-54.

Gorman, C. M., Moffat, L. F. and Howard, B. H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase inmammalian cells. Mol. Cell. Biol. **2**; 1044-1051.

Gottesman, M. E. and Weisberg, R. A. (1971). Prophage insertion and excision in the bacteriophage lambda, A. D. Hershley ed. (Cold Spring Harbour New York: Cold Spring Harbour Laboratory), pp 113-138.

Graham, F. L. and van der Eb, A. J. (1973). A new technique for the assay of infectivity of human Adenovirus 5 DNA. Virology **52**; 456-467.

Granados, R. R. and Williams, K. A. (1986). In vivo infection and replication of baculoviruses. In the Biology of Baculoviruses, Vol 1, eds. Granados, R. and Federici, B. Boca Raton, Fla: CRC Press. pp 89-108.

Granados, R. R. and Williams, K. A. (1986). In vivo infection and replication of baculoviruses. In the Biology of Baculoviruses, Vol 1, eds Granados, R. and Federici, B. Boca Raton, Fla: CRC Press. pp 90-127.

Granger-Schnarr, M., Lloubes, R., De Murcia, G. and Schnarr, M. (1988). Specific Protein-DNA complexes: Immunodetection of the Protein Component after Gel Electrophoresis and Western Blotting. Anal. Biochem. **174**; 235-238.

Grosveld, F., van Assendelft, G. B., Greaves, D. R. and Kollias, G. (1987). Position-Independent, high-level Expression of the Human β-Globin Gene in Transgenic Mice. Cell **51**; 975-985.

Griffiths, C. M. (1994). Baculovirus expression vectors: advances and applications. Exp. Opin. Ther. Patents **4**; 1065-1082.

Gruss, B. and Khoury, G. (1980). Rescue of a splicing defective mutant by insertion of an heterologous intron. Nature **286**; 634-637.

Gu, H., Zou, Y. R., and Rajewsky K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP mediated gene targeting. Cell **73**; 1155-1164.

Gu, H., Math, J. D., Orban, P. C., Mossmann, H. and Rajewsky, K. (1994). Deletion of a DNA polymerase β gene segment in T cells using Cell Type-specific Gene Targeting. Science **265**; 103-106.

Guarino, L. A. and Summers, M. D. (1986). Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene. J. Virol. **57**; 563-571.

Guarino, L. A. and Summers, M. D. (1987). Nucleotide sequence and temporal expression of a baculovirus regulatory gene. J. Virol. **61**; 2091-2099.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**; 557-580.

Haring, M. A., Rommens, C. M. T., Nijkamp, H. J. J. and Hille, J. (1991). The use of transgenic plants to understand transposition mechanisms and to develop transposon tagging strategies. Plant Mol. Biol. **16**; 449-461.

Harrap, K. (1972a). The structure of nuclear polyhedrosis virus in the inclusion body. Virology **50**; 114-123.

Harrap, K. and Longworth, J. (1974). An evaluation of purification methods for baculoviruses. J. Invertbr. Pathol. **24**; 55-62.

Hasnain, S. E., Nakhai, B., Ehtesham, N. Z., Sridhar, P., Ranjan, A., Talwar, G. P. and Jha, P. K. (1994). β-subunit of human chorionic gonadotrophin hormone and firefly

luciferase simultaneously synthesised in insect cells using a recombinant baculovirus are differentially expressed and transported. DNA and Cell Biol. **13**; 274-282.

Hasty, P., Ramirez-Solis, R., Krumlauf, R. and Bradley, A. (1991). Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. Nature **350**; 243-246.

Hattman, S., Brooks, J. E.and Masurekar, M. (1978). Sequence specificity of the P1 modification methylase (M. *Eco*P1) and the DNA methylase (M. *Eco*dam) controlled by the Escherichia coli dam gene. J. Mol. Biol. **126**; 367-380.

Heartlein, M. W., Knoll, J. H. M. and Latt, S. A. (1988). Chromosome Instability Associated with Human Alphoid DNA Transfected into the Chinese Hamster Genome. Mol. and Cell. Biology **8**; 3611-3618.

Henrik, A., Dale, E. C., Lee, E. and Ow, D. W. (1995). Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. The Plant Journal 7; 649-659.

Hildebrandt, E. R. and Cozzarelli, N. R. (1995). Comparison of Recombination in Vitro and in E. coli Cells: Measure of the Effective Concentration of DNA In Vivo. Cell **81**; 331-340.

Hill-Perkin, M. S. and Possee, R. D. (1990). A baculovirus expression vector derived from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus. J. Gen. Virol. **71**; 971-976.

Hippenmeyer, P. and Highkin, M. (1993). High level, stable production of recombinant proteins in mammalian cell culture using the Heresvirus VP16 Transactivator. Bio/Technology **11**; 1037-1041.

Hochuli, E., Dobeli, H. and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J. Chromatogr. **411**; 177-184.

Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Stubber, D. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/Technology **6**; 1321-1325.

Hoess, R., Ziese, M. and Sternberg, N. (1982). P1 site-specific recombination: nucleotide sequence of the recombining sites. Proc. Natl. Acad. Sci. USA. **79**; 3398-3402.

Hoess, R. H., Wierzbicki, A. and Abremski, K. (1984). The role of the loxP spacer region in P1 site-specific recombination. Nucleic Acids Research **14**; 2287-2300.

Hoess, R. H. and Abremski, K. (1984). Interaction of bacteriophage P1 recombinase with the recombining site loxP. Proc. Natl. Acad. Sci. USA **81**; 1026-1029.

Hoess, R. and Abremski, K. (1985). Mechanism of strand cleavage and exchange in the Cre-loxP site-specific recombination system. J. Mol. Biol. **181**; 351-362.

Hoess, R., Wierzbicki, A. and Abremski, K. (1987). Isolation and characterisation of intermediates in site-specific recombination. Proc. Natl. Acad. Sci. USA. **84**; 6840-6844.

Holliday, R. (1964). A mechanism for gene conversion in fungi. Genet. Res. **5**; 282-304.

Hoss, A., Moarefi, I., Scheidtmann, K. H., Cisek, L. J., Corden, J. L., Dornreiter, I., Arthur, A. K. and Fanning, E. (1990). Altered phosphorylation pattern of simian virus 40 T antigen expressed in insect cells by using a baculovirus vector. J. Virol, **64**; 4799-4807.

Holt, L. C. and May, G. S. (1993). A novel phage λ replacement Cre-loxP vector that has been automatic subcloning capabilities. Gene **133**; 95-97.

Ikeda, H. and Tomizawa, J. I. (1968). Prophage P1, an extrachromosomal element. Cold spring harbour symp. Quant. Biol. **33**; 791-798.

Ikeda, H., Trowsdale, J. and Saito, I. (1988). Mulcos: a vector for amplification and simultaneous expression of two foreign genes in mammalian cells. Gene **71**; 19-27.

Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F. and Boyer, H. W. (1977). Expression in *Escherichia coli* of a chemically synthesised gene for human somatostatin. Science. **198**; 1056-1063.

Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A., Nordheim, A. and Stunnenberg, H. G. (1991). Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. Proc. Natl. Acad. Sci. USA. **88**; 8972-8976.

Jarvis, D. L. and Summers, M. D. (1989). Glycosylation and secretion of human tissue plasminogen activatior in recombinant baculovirus-infected insect cells. Mol. Cell. Biol. **9**; 214-223.

Jarvis, D. L., Fleming , J. G. W., Kovacs, G. R., Summers, M. D. and Guarino, L. A. (1990). Use of early baculovirus promoters for continuous expression and efficient processing of foreign gene products in stably transformed lipidopteran cells. Bio/Technology. **8**; 950-955.

Jarvis, D. L., Summers, M. D., Garcia, A. and Bohlmeyer, D. A. (1993). Influence of different signal peptides and prosequences on expression and secretion of human tissue plasminogen activator in the baculovirus system. J. Biol. Chem. **268**; 16754-16762.

Jayaram, M. (1994). Phosphoryl transfer in Flp recombination: a template for strand transfer mechanisms. TIBS **19**; 78-82.

Jimenez, A. and Davies, J. (1980). Expression of a transposable antibiotic resistance element in *Saccharomyces cerevisiae*. Nature **287**; 869-871.

Johal, G. S. and Briggs, S. P. (1992). Reductase activity encoded by the HM1 disease resistance gene in maize. Science **258**; 985-987.

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.

Kaestner, K. H., Montaliu, L., Kern, H., Thulke, M. and Schütz, G. (1994). Universal β-galactosidase cloning vectors for promoter analysis and gene targeting. Gene **148**; 67-70.

Kalderon, D., Roberts, B. L., Richardson, W. D., and Smith, A. E. (1984). A short amino acid sequence able to specify nuclear location. Cell **39**; 499-509.

Kashanchi, F., Duvall, J. F. and Brady, J. N. (1992). Electroporation of viral transactivator proteins into lymphocyte suspension cells. Nucl. Acids Res. **20**; 4673-4674.

Kaufman, R. J. and Sharp, P. A. (1982). Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase cDNA gene. J. Mol. Biol. **159**; 601-621.

Kaufman, R. J., Wasley, L. C., Spilotes, A. J., Gossels, S. D., Latt, S. A., Larsen, G. R. and Kay, R. M. (1985). Co-amplification and co-expression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in chinese hamster ovary cells. Mol. Cell. Biol. **5**; 1750-1759.

Kaufman, R. J., Murtha, P. and Davies, M. V. (1987). Translation efficiency of polycistronic mRNAs and their utilisation to express heterologous genes in mammalian cells. EMBO J. **6**; 187-193.

Kaufman, R. J. (1990). Vectors used for expression in mammalian cells. Methods in Enzymology **185**; 487-511.

Keddie, B. A., Aponte, G. W. and Volkman, L. E. (1989). The pathway of infection of *Autographa californica* nuclear polyhedrosis virus in an insect host. Science **243**; 1728-1730.

Kinosita, K. Jr. and Tsong, T. Y. (1977). Voltage-induced pore formation and hemolysis of human erythrocytes. Biochim. Biophys. Acta **471**; 227-242.

Kelly, D. C. (1982). Baculovirus replication. J. Gen. Virol. **63**; 1-13.

Kitts, P. 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.

Kitts, P. A., Ayres, M. D. and Possee, R. D. (1990). Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. Nucl. Acids. Res. **18**; 5667-5672.

Kitts, P. A. and Possee, R. D. (1993). A method for producing recombinant baculovirus expression vectors at high frequency. BioTechniques **14**; 810-817.

Klehr, D., Maass, K. and Bode, J. (1991). Scaffold-attached human interferon-beta domain can be used to enhance the stable expression of genes under the control of various promoters. Biochemistry **30**; 1264-1270.

Koller, B. H. and Smithies, O. (1992). Altering genes in animals by gene targeting. Annu. Rev. Immunol. **10**; 705-730.

Kozak, M. (1987). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J. Mol. Biol. **196**; 947-950.

Kozak, M. (1989). The scanning model for translation: an update. J. Cell. Biol. **108**; 229-241.

Krasnow, M. A. and Cozzarelli, N. R. (1982). Catenation of DNA rings by topoisomerases-Mechanism of control by spermidine. J. Biol. Chem. **257**; 2687-2693.

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.

Kriegler, M. (1990). Gene transfer and expression. A laboratory manual. Stockton Press, New York.

Kühn, R., Schwenk, F., Aguet, M. and Rajewsky, K. (1995). Inducible Gene Targeting in Mice. Science **269**; 1427-1429.

Kuzio, J., Rohel, D. Z., Curry, C. J., Krebs, A., Carstens, E. B. and Faulkner, P. (1984). Nucleotide sequence of the p10 polypeptide gene of *Autographa californica* nuclear polyhedrosis virus. Virology **139**; 414-418.

Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D. and Constantini, F. D. (1983). A foreign beta-Globin Gene In Transgenic Mice-Integration at abnormal chromosomal positions and expression in inappropraite tissue. Cell **34**; 343-358.

Laemmili, U. K. (1970). Cleavage of Structural proteins during the assembly of the head of bacteriophage T4. Nature **227**; 680-685.

Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R. and Berns, A. (1991). Simplified mammalian DNA isolation procedure. Nucl. Acids Res. **19**; 4293.

Lake, R. A. and Owen, M. J. (1991). Chapters 16-18 in Gene Transfer and Expression Protocols, Methods in Molecular Biology, Vol 7; Murray E. J. (Ed.) Humana.

Lakso, D., Cavenee, W. and Nordenskjold, M. A. (1991). Loss of constitutional heterozygosity in human cancer. Ann. Rev. Genet. **25**; 281-314.

Lakso, M., Sauer, B., Mosinger, B. J., Lee, E. J., Manning, R. W., Yu, S. H., Mulder, K. L. and Westphal, H. (1992). Targeted oncogene activation by site-specific recombination in transgenic mice. Proc. Natl. Acad. Sci. USA **89**; 6232-6236.

Lederberg, S. (1957). Suppression of the multiplication of heterologous bacteriophages in lysogenic bacteria. Virology **3**; 496-513.

Lee, H. H. and Miller, L. K. (1978). Isolation of genotype variants of Autographa Californica nuclear polyhedrosis virus. J. Virol **27**; 754-764.

Lennox, E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. Virology **1**; 190-206.

Leong, J. M., Nunes-Dubly, S. E., Lesser, C. F., Youderian, P., Sussland, M. M. and Landy, A. (1985a). The PHI-80 and P22 attachment sites-Primary structure and interaction with *Escherichia coli* integration host factor. J. Biol. Chem. **260**; 4468-4477.

Levinson, A. D. (1990). Expression of heterologous genes in mammalian cells. In Methods in Enzymology (Academic Press, New York) **185**; 485-527.

Lim, K. and Chae, C. -B. (1989). A simple assay for DNA transfection by incubation of cells in culture dishes with substrates for βeta-galactosidase. BioTechniques **7**; 576-579.

Lin, S. -Y. and Riggs, A. D. (1975). The general affinity of *lac* repressor for *E. coli* DNA: implications for gene regulation in prokaryotes and eukaryotes. Cell **4**; 107-111.

Lindahl, G. and Sunshine, M. (1972). Excision -deficient mutants of bacteriophage P2. Virology **49**; 180-187.

Longworth, J. L. (1983) Current problems in insect virus taxonomy: in a critical appraisal of virus taxonomy. Matthews, R.E.F., Ed., CRC Press, Boca Raton, Fla; pp 123.

Lubiniecki, A., Arathoon, R., Polastri, G., Thomas, J., Wiebe, M., Garnick, R., Jones, A., van Reis, R. and Builder, S. (1989). Selected strategies for manufacture and control of recombinant tissue plasminogen activator prepared from cell culture. In: Advances in Animal Cell Biology and Technology for Bioprocesses. Spier, R. E., Griffiths, J. B., Stephens, J., Crooy, P. J. eds. pp 442-449.

Luckow, V. A. and Summers, M. D. (1988). Trends in the development of baculovirus expression vectors. Bio/Technology **6**; 47-55.

Luckow, V. A. and Summers, M. D. (1989). High level expression of non-fused foreign genes with *Autographa californica* nuclear polyhedrosis virus expression vector. Virology **170**; 31-39.

MacDonald, C. (1991). The expression of foreign genes in animal cells. Chemistry and Industry **18**; 668-673.

Mack, A., Sauer, B., Abremski, K. and Hoess, R. (1992). Stoichiometry of the Cre recombinase bound to the loxP recombining site. Nucleic Acids Research **17**; 4451-4455.

Maeda, S., Kawai, T., Obinata, M., Fujiwara, H., Horiuchi, T., Saeki, Y., Sato, Y. and Furusawa, M. (1985). Production of human alpha-interferon in silkworms using a baculovirus vector. Nature **315**; 592-594.

Maeda, S. (1989). Expression of foreign genes in insects using baculovirus vectors. Annu. Rev. Entamol. **34**; 351-372.

Makkuni, J. (1994). Phosphoryl transfer in flp recombination: a template for strand transfer mechanism. TIBS. **19**; 78-82.

Maniatis, T. Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning ; A laboratory manual. (Cold Spring Harbour Laboratory, Cold Spring Harbour, N. Y.).

Maruyama, I. N. and Brenner, S. (1992). A selective λ phage cloning vector with automatic excision of the insert in a plasmid. Gene **120**; 135-141.

Matsuura, Y., Possee, R. D., Overton, H. A., and Bishop, D. H. L. (1987). Baculovirus expression vectors: The requirements for high level expression of proteins, including glycoproteins. J. Gen. Virol. **67**; 1515-1529.

McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J. (1990). Phage Antibodies-Filamentous phage displaying antibody variable domains. Nature **348**; 552-554.

McComb, R. B. and Bowers, G. N. Jr. (1972). Study of optimum buffer conditions for measuring alkaline phosphatase activity in human serum. Clin. Chem. **18**; 97-104.

McCutchan, J. H. and Pagano, J. S. (1968). Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. J. Natl. Cancer Inst. **41**; 351-357.

McLeod, M., Volkert, F. and Broach, J. (1984). Cold Spring Harbour Symp. Quant. Biol., **49**; 779-787.

McMahon, A. P. and Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. Cell **62**; 1073-1085.

Meese, E., Olson, S., Leis, L. and Trent, J. (1990). Quick method for high yields of λ bacteriophage DNA. Nuc. Acids Res. **18**; 1923.

Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981). Identification of DNA sequences required for the transcription of the human α1-globin gene in a new SV40 host-vector system. Cell **27**; 279-282.

Melton, D. W. (1994). Gene targeting in the mouse. Bioessays 16; 633-638.

Miller, L. K. (1981). A virus vector for genetic engineering in invertebrates. In Genetic Engineering in the Plant Sciences. ed. Panopoulos, N., New York. pp 203-224.

Miller, L. K., Lingg, A. J. and Bulla Jr., L. A. (1983a). Bacterial viral and fungal insecticides. Science **219**; 715-721.

Miller, L. K., Miller, D. and Adany, M. (1983b). An insect virus for genetic engineering: developing baculovirus polyhedrin substitution vectors. In Genetic Engineering In Eukaryotes, eds. Lurquin, P. and Kleinhofs, A. New York: Plenum pp 89-97.

Miller, L. K. (1988a). Baculoviruses as gene expression vectors. Annu. Rev. Microbiol. **42**; 131-137.

Miller, L. K. (1988b). Baculovirus as gene expression vectors. Annu. Rev. Microbiol. **42**; 177-199.

Miller, A. D. and Rosman, G. J. (1989). Improved Retroviral Vectors for Gene Transfer and Expression. BioTechniques **7**; 980-990.

Mirkovitch, J., Mirault, M., -E. and Laemmli, U. K. (1984). Organisation of the higher-order chromatin loop-specific DNA attachment sites on nuclear scaffold. Cell **39**; 223-232.

Miyamoto, C., Smith, G. E., Farrell-Towt, J., Chizzonite, R., Summers, M. D. and Ju, G. (1985). Production of c-myc protein in insect cells infected with a baculovirus expression vector. Mol. Cell. Biol, **5**; 2860-2865.

Mulligan, R. C. and Berg, P. (1980). Expression of a bacterial gene in mammalian cells. Science **209**; 1422-1427.

Murphy, A. J. M. and Schimke, R. T. (1991). pλzd39: a new type of cDNA expression vector for low background, high efficiency directional cloning. Nucl. Acids Res. **19**; 3403-3408.

Needham, M., Egerton, M., Millest, A., Evans, S., Popplewell, M., Cerillo, G., McPheat, J., Monk, A., Jack, A., Johnstone, D. and Hollis, M. (1995). Further development of the locus control region/murine erythroleukemia expression system:

high level expression and characterisation of recombinant human calcitonin receptor. Prot. Exp. and Purific. **6**; 124-131.

Niwa, H., Yamamura, K. and Miyazaki, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene **108**; 193-200.

Oeltinger, M. A., Schatz, D. G., Gorka, C. and Baltimore, D. (1990). Rag-1 and Rag-2, adjacent genes that synergistically activate V(D)J recombination. Science **248**; 1517-1523.

O'Gorman, S., Fox, D. T. and Wahl, G. M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science **251**; 1351-1355.

O'Reilly, D. R. and Miller, L. K. (1989). A baculovirus blocks insect molting by producing ecdysteriod UDP-glycosyl transferase. Science **245**; 1110-1112.

O'Reilly, D. R., Passarelli, A. L., Goldman, I. F. and Miller, L. K. (1990). Characterisation of the DA26 gene in a hypervariable region of the *Autographa californica* nuclear polyhedrosis virus genome. J. Gen. Virol. **71**; 1029-1037.

O'Reilly, D. R., Brown, M. R. and Miller, L. K. (1992). Alteration of ecdysteroid metabolism due to baculovirus infection of the fall armyworm *Spodoptera frugiperda*: host ecdysteroids are conjugated with galactose. Insect Biochem. Mol. Biol. In press.

Onouchi, H., Yokoi, K., Machida, C., Matsuzaki, H., Oshima, Y., Matsuoka, K., Nakumura, K. and Machida, Y. (1992). Operation of an efficient system of *Zygosaccharomyces-Rouxii* in Tobacco Cells. Nucl. Acids Res. **19**; 6373-6378.

Page, M. J. (1989). p36c: an improved expression vector for producing high levels of mature recombinant proteins. Nucl. Acids. Res. **17**; 454.

Page, M. J. and Sydenham, M. A. (1991). High level expression of the humanised monoclonal antibody CAMPATH-1H in chinese hamster ovary cells. Bio/Technology. **9**; 64-68.

Palazzolo, M. J., Hamilton, B. A., Ding, D., Martin, C. H., Mead, D. A., Mierendorf, R. C., Ragharan, K. V., Meyerowitz, E. M. and Lipshitz, H. D. (1990). Phage λ cDNA cloning vectors for subtractive hybridisation, fusion-protein synthesis and Cre-loxP automatic plasmid subcloning. Gene **88**; 25-36.

Pallavacini, M. G., DeTeresa, P. S., Rosette, C., Gray, J. and Wurm, F. M. (1990). Effects of methotrexate on transfected DNA stability in mammalian cells. Mol. Cell. Biol. **10**; 401-404.

Palmer, T. D., Hock, R. A., Osborne, W. R. A. and Miller, A. D. (1987). Efficient retrovirus-mediated transfer and expression of a human adenosine deaminase gene in diploid skin fibroblasts from an adenosine deaminase-deficient human. Proc. Natl. Acad. Sci. USA **84**; 1055-1059.

Palmiter, R. D. and Brinster, R. L. (1986). Germ-line transformation of mice. Ann. Rev. Genet. **20**; 465-499.

Patel, G., Nasmyth, K. and Jones, N. (1992). A new method for the isolation of recombinant baculovirus. Nucl. Acids. Res. **20**; 97-104.

Peakman, T., Page, M. and Gewert, D. (1989). Increased recombinational efficiency in insect cells irradiated with short wavelength ultra-violet light. Nucl. Acids. Res. **13**; 5403.

Peakman, T. C., Harris, R. A. and Gewert, D. R. (1992a). Highly efficient generation of recombinant baculoviruses by enzymatically mediated site-specific in-vitro recombination. Nucleic Acids Research **20**; 495-500.

Peakman, T. C., Charles, I. G., Sydenham, M. A., Gewert, D. R., Page, M. J. and Mackoff, A. J. (1992b). Enhanced expression of recombinant proteins in insect cells using a baculovirus vector containing a bacterial leader sequence. Nucl. Acids. Res. **20**; 6111-6112.

Peakman, T. C., Worden, J., Harris, R. H., Cooper, H., Tite, J., Page, M. J., Gewert, D. R., Bartholemew, M., Crowe, J. S. and Brett, S. (1994). Comparison of expression of a humanised monoclonal antibody in mouse NSO myeloma cells and Chinese Hamster Ovary cells. Human Antibodies and Hybridomas **5**; 65-74.

Pennock, G. D., Shoemaker, C. and Miller, L. K. (1984). Strong and regulated expression of Escherichia coli β-galactosidase in insect cells with a baculovirus vector. Mol. Cell. Biol. **4**; 399-406.

Peters, R. (1983). Nuclear-envelope permeability measured by fluorescence microphotolysis of single liver-cell nuclei. J. Biol. Chem. **258**; 11427-11429.

Phi-Van, L., von Kries, J. P., Ostertag, W. and Stratling, W. H. (1990). The chicken lysozyme 5' Matrix Attachment Region Increases Transcription from a Heterologous Promoter in Heterologous Cells and Dampens Position Effects on the Expression of Transfected Genes. Mol. and Cell. Biology **10**; 2302-2307.

Pillot, T., Barbier, A., Visvikis, A. and Lazac'h, K. (1996). Single-Step Purification of Two functional Human Apolipoprotein E Variants Hyperexpressed in *Escherichia coli*. Protein Expression and Purification. 7; 407-414.

Poljak, L., Seum, C., Mattioni, T. and Laemmli, U. K. (1994). SARs stimulate but do not confer position independent gene expression. Nucl. Acids Res. **22**; 4386-4394.

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. Nature **258**; 598-599.

Porter, A. C. G. (1989). Designer genomes. Technique **1**; 53-65.

Possee, R. D. (1986). Cell-surface expression of influenza virus haemagglutinin in insect cells using a baculovirus vector. Virus Res. **5**; 43-59.

Qin, M., Bayley, C., Stockton, T. and Ow, D. W. (1994). Cre recombinase-mediated site-specific recombination between plant chromosome. Proc. Natl. Acad. Sci. USA. **91**; 1706-1710.

Rabbitts, T. H. (1994). Chromosomal translocation in human cancer. Nature **372**; 143-149.

Ramirez-Solis, R., Liu, P. and Bradley, A. (1995). Chromosome engineering in mice. Nature **378**; 720-724.

Rankl, N. B., Rice, J. W., Gurganus, T. M., Barbee, J. L. and Burns, D. J. (1994). The production of an active protein kinase c-δ in insect cells is greatly enhanced by the use of the basic protein promoter. Protein Expression and Engineering **5**; 346-356.

Ray, J. and Gage, F. H. (1992). Gene transfer into established and primary fibroblast cell lines: Comparison of transfection methods and promoters. BioTechniques **13**; 598-603.

Reff, M. E. and Pfarr, D. S. (1992). Production of proteins by mammalian cells using dominant selectable and amplifiable markers. *In Gene Amplification in* Mammalian cells. Edited by Kellems, R. E. New York; Marcel Dekker, Inc. pp 355-367.

Reiss, M., Jastreboff, M. M., Bertinol, H. R. and Narayam, R. (1986). DNA-mediated gene transfer into epidermal cells using electroporation. Biochem. Biophys. res. Commun. **137**; 244-249.

Rinchik, E. M. and Russell, L. B. (1990). Genome Analysis Volume II, Genetic and physical mapping (eds Davies, K. E. and Tilgham, S. M.; Cold Spring Harbour Lab. Press, New York). pp 121-158.

Roberts, J. M. and Axel, R. (1982). Amplification and correction of transformed genes. In Gene Amplification. ed. Scinke, R. T. Cold Spring Harbour Press, New York. pp 152-155.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**; 487-491.

Sambrook, J., Fritch, E. F. and Maniatis, T. (1989). Molecular Cloning. A Laboratory Manual, 2nd. ed. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N. Y.

Sandmeier, H., Iida, S., Meyer, J., Hiestand-Nauer, R. and Arber, W. (1990). Site-specific DNA recombination system. Min of plasmid P15B-A cluster of overlapping invertible DNA segments. Proc. Natl. Acad. Sci. USA **87**; 1109-1113.

Sanger, F., Nicklen, S. and Coulsen, A. R. (1977). DNA sequencing and chain-termination inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74; 5463-5467.

Sauer, B. (1987). Functional expression of the Cre-lox site-specific recombination system in the yeast *Saccaromyces cerevisiae*. Molecular and Cellular Biology 7; 2087-2096.

Sauer, B. and Henderson, N. (1987). Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. Nucleic Acid Research 17; 147-161.

Sauer, B. and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc. Natl. Acad. Sci. USA. 85; 5166-5170.

Sauer, B. and Henderson, N. (1989). Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. Nucl. Acids Res. 17; 147-161.

Sauer, B. and Henderson, N. (1990). Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. The New Biologist 2; 441-449.

Sauer, B. (1992). Identification of cryptic lox sites in the yeast genome by selection for Cre-mediated chromosome translocations that confer multiple drug resistance. J. Mol. Biol. 223; 911-928.

Sauer, B. (1993). Manipulation of Transgenes by Site-Specific Recombination: Use of Cre Recombinase. Methods Enzymology 225; 890-900.

Sauer, B. (1994). Recycling selectable markers in yeast. BioTechniques 16; 1086-1088.

Schatz, D. G., Oettinger, M. A. and Baltimire, D. (1989). The V(D)J recombination activating gene, Rag-1. Cell 59; 1035-1048.

Schulz, D. W., Taylor, A. F. and Smith, G. R. (1983). *Escherichia coli* RecBC pseudorevertants lacking chi recombinational hotspot activity. J. Bacteriol 155; 664-680.

Schwartzberg, P. I., Goff, S. P. and Robertson, E. J. (1989). Germ-line transmission of a c-Abl mutation produced by targeted gene disruption in ES cells. Science 246; 799-803.

Scott, J. R. (1968). Clear plaque mutants of phage P1. Virology 41; 66-71.

Scott, J. R. and Kropf, M. M. (1977). Location of new clear plaque genes on the P1 map. Virology 82; 362-368.

Searle, P. F., Stuart, G. F. and Palmiter, R. D. (1985). Building a metal-responsive promoter with synthetic regulatory elements. Mol. Cell. Biol. 5; 1480-1489.

Segev, N., Laub, A. and Cohen, G. (1980). A circular form of bacteriophage P1 DNA made in lytically infected cells of *Escherichia coli*. 1. Characterisation and kinetics of formation. Virology **101**; 261-271.

Segev, N. and Cohen, G. (1981). Control of circularisation of bacteriophage P1 DNA in Escherichia coli. Virology **114**; 333-342.

Senecoff, J. F., Rossmeisl, P. J. and Cox, M. M. (1988). DNA recognition by the FLP recombinase of the yeast 2μ plasmid: a mutational analysis of the FLP binding site. J. Mol. Biol. 201; 405-421.

Short, J. M., Fernandez, J. M., Sorge, J. A. and Huse, W. D. (1988). λZAP: a bacteriophage λ expression vector with in vivo excision properties. Nucl. Acids Res. **16**; 7583-7600.

Smith, G. E., Vlak, J. M. and Summers, M. D. (1982). In-vitro translation of Autographa Californica nuclear polyhedrosis virus early and late mRNAs. J. Virol **44**; 199-208.

Smith, G. R. (1983). Chi hotspots of generalised recombination. Cell **34**; 709-710.

Smith, G. E., Summers, M. D. and Frazer, M. J. (1983a). Production of human beta interferon in insect cells infected with baculovirus expression vector. Mol. Cell. Biol. **3**; 2156-2165.

Smith, G. E., Vlak, J. M. and Summers, M. D. (1983b). Physical analysis of Autographa californica nuclear polyhedrosis virus transcripts for polyhedrin and 10,000 molecular-weight protein . J. Virol **45**; 215-225.

Smith, G. E., Fraser, M. J. and Summers, M. D. (1983c). Molecular engineering of the Autographa Californica nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene. J. Virol **46**; 584-593.

Smith, A. J. H., De Sousa, M., Kwabi-Addo, B., Happell-Parton, A., Impey, H. and Rabbitts, P. (1995). A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. Nature Genetics **9**; 376-385.

Soriano, P., Montgomery, C., Geske, R. and Bradley, A. (1991). Targeted disruption of the c-src proto-oncogene leads to osteopetrosis in mice. Cell **64**; 693-702.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**; 503-517.

Southern, P. J. and Berg, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Gen. **1**; 327-341.

Sridhar, P., Panda, A. K., Pal, R., Talwar, G. P. and Hasnian, S. E. (1993). Temporal nature of the promoter and not relative strength determines the expression of an extensively processed protein in a baculovirus system. FEBS Lett. **315**; 282-286.

Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A. and Melton, D. W. (1994). Use of double-replacement gene targeting to replace the murine α-lactalbumin gene with its human counterpart in embryonic stem cells and mice. Mol Cell. Biol. **14**; 1009-1016.

Stark, G. R. and Wahl, G. M. (1984). Gene Amplification. Ann. Rev. Biochem. **53**; 447-491.

Stark, G. R., Debatisse, M., Guilotto, E. and Wahl, G. M. (1989). Recent progress in understanding mechanisms of mammalian DNA amplification. Cell **57**; 901-908.

Stark, W. M., Boocock, M. R. and Sherratt, D. J. (1992). Catalysis by site-specific recombinases. Trends Genet. **8**; 432-439.

Steinmetz, M., Vematsu, Y. and Lindahl, K. F. (1987). Hotspots of homologous recombination in mammalian genomes. TIG **3**; 7-10.

Sternberg, N. (1978). Demonstration and analysis of P1 site-specific recombination using λ-P1 hybrid phages constructed *in vitro*. Cold Spring Harbour Symp. Quant. Biol. **43**; 1143-1146.

Sternberg, N. (1979). A characterisation of bacteriophage P1 DNA fragments cloned in a lambda vector. Virology **96**; 129-142.

Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M. and Hoess, R. (1980). Site-specific recombination and its role in the life cycle of bacteriophage P1. Cold Spring Harbour Symp. Quant. Biol **45**; 297-309.

Sternberg, N. and Hamilton, D. (1981). Bacteriophage P1 site-specific recombination I. Recombination between loxP sites . J. Mol. Biol. **150**; 467-486.

Sternberg, N., Hamilton, D. and Hoess, R. (1981). Bacteriophage P1 site-specific recombination II. Recombination between loxP and the bacterial chromosome. J. Mol. Biol. **150**; 487-507.

Sternberg, N., Sauer, B., Hoess, R. and Abremski, K. (1986). Bacteriophage P1 Cre gene and its regulatory region. Evidence for multiple promoters and fro regulation by DNA methylation. J. Mol. Biol. **187**;197-212.

Stief, A. D., Winter, M., Stratling, W. H. and Sippel, A. E. (1989). A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature **341**; 343-345.

Stigbrand, T. (1984). Present states and future trends of human alkaline phosphatases. Prog. Clin. Biol. Res. **166**; 3-14.

Sugden, B., March, K. and Yates, J. (1985). A vector that replicates as a plasmid and can be efficiently selected in B-lymphoblasts transformed by Epstein-Barr virus. Mol. Cell. Biol. **5**; 410-413.

Summers, M. and Anderson, D. (1972). Granulosis virus deoxyribonucleic acid: a closed double-stranded molecule. J. Virol. **9**; 710-713.

Summers, M. D. and Smith G. E. (1987). A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agric. Exp. Stn. Bull. **1555**; 1-56.

Swanson, J. A., Lee, M. and Knapp, P. E. (1991). Cellular dimensions affecting the nucleocytoplasmic volume ratio. J. Cell Biol. **115**; 941-948.

Takeshita, S., Tezuka, K., Takahashi, M., Honkawa, H., Matsuo, A., Matsuishi, T. and Hashimoto-Gotoh, T. (1988). Tandem gene amplification in vitro for rapid and efficient expression of animal cells. Gene **71**; 9-18.

Tessier, D. C., Thomas, D. Y., Khouri, H. E., Laliberte, F. and Vernet, T. (1991). Enhanced secretion from insect cells of a foreign protein fused to the honeybee melittin signal peptide. Gene. **98**; 177-183.

Thiem, S. M. and Miller, L. K. (1989a). Identification, sequence and transcriptional mapping of the major capsid protein gene of the baculovirus *Autographa californica* nuclear polyhedrosis virus. J. Virol. **63**; 2008-2018.

Theim, S. M. and Miller, L. K. (1989b). A baculovirus gene with a novel transcription pattern encodes a polypeptide with a zinc finger and a leucine zipper. J. Virol. **63**; 4489-4497.

Theim, S. M. and Miller, L. K. (1990). Differential gene expression mediated by late, very late, and hybrid baculovirus promoters. Gene **91**; 87-94.

Urlaub, G. and Chasin, L. (1980). Isolation of Chinese hamster ovary cell mutants deficient in dihydrofolate reductase activity. Proc. Natl. Acad. Sci **77**; 4216-4220.

Vaeck, M., Reynaerts, A., Höfte, H., Jansens, S., De Beuck-leer, M., Dean, C., Zabeau, M., Van Montagu, M. and Leemans, J. (1987). Transgenic plants protected from insect attack. Nature **328**; 33-37.

Van der Beek, C. P., Saaijer-Riep, J. D. and Vlak, J. M. (1980). On the origin of the polyhedrin protein of Autographa californica nuclear polyhedrosis virus. Virology **100**; 326-333.

Van Haaren, M. J. J. and Ow, D. W. (1993). Prospects of applying a combination of DNA transposition and site-specific recombination in plants: a strategy for gene identification and cloning. Plant Molecular Biology **23**; 487-491.

Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. and McCawley, P. (1977). The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera Noctuidae). In Vitro **13**; 213-217.

Vialard, J., Lalumière, M., Vernet, T., Briedis, D., Alkhatils, G., Henning, D., Levine, D. and Richardson, C. (1990). Synthesis of the membrane fusion and hemagglutinin proteins of measles virus, using a novel baculovirus vector containing the β-galactosidase gene. J. Virol. **64**; 37-50

Vlak, J. M., Schouten, A., Usmany, M., Belsham, G. J., Klinge-Roode, E. C., Maule, A. J., Van Lent, J. W. M. and Zuidema, D. (1990). Expression of Cauliflower mosaic virus gene I using a baculovirus vector based upon the p10 gene and a novel selection method. Virology **179**; 312-320.

Vologodskii, A. V., Levene, S. D., Klenin, K. V., Frank-Kamenestskii, M. D. and Cozzarelli, N. R. (1992). Conformational and thermodynamic properties of supercoiled DNA. J. Mol. Biol. **227**; 216-218.

von Hippel, P. H. and Berg, O. G. (1986). On the specificity of DNA-protein interactions. Proc. Natl. Acad. Sci. USA. **83**; 1608-1612.

Walker, D. H. Jr. and Walker, J. T. (1975). Genetic studies of coliphage P1. 1. Mapping by use of prophage deletions. J. Virol. **16**; 525-534.

Walker, D. H. Jr. and Walker, J. T. (1976). Genetic studies of coliphage P1. 1. Mapping by use of prophage deletions. J. Virol **20**; 177-187.

Wang, X., Guat Ooi, B. and Miller, L. K. (1991). Baculovirus vectors for multiple gene expression and for occluded virus production. Gene **100**; 131-137.

Waterhouse, P., Griffiths, A. D., Johnson, K. S. and Winter, G. (1993). Combinatorial Infection and in vivo recombination: a strategy for making large phage antibody repertoires. Nucl. Acids Res. **21**; 2265-2266.

Watson, J. D. and Crick, F. H. C. (1953). Molecular structure of nucleic acid. A structure for deoxyribose nucleic acid. Nature **171**; 737-738.

Wenger, R. H., Moreau, H. and Neilsen, P. J. (1994). A comparison of different promoter, enhancer, and cell type combinations in transient transfections. Anal. Biochem. **221**; 416-418.

Weidle, U. H., Buckel, P. and Wienberg, J. (1988). Amplified expression constructs for human tissue-type plasminogen activator in Chinese hamster ovary cells: instability in the absence of selective pressure. Gene **66**; 193-203.

Weisberg, B. A. (1983). Role for DNA homology in Site-Specific Recombination. The isolation and characterisation of a site affinity mutant of coliphage λ. J. Mol. Biol. **170**; 319-342.

Whittle, N., Adair, J., Lloyd, C., Jenkins, L., Devine, J., Schlom, J., Raubitschek, A., Colcher, D. and Bodmer, M. (1987). Expression in COS cells of a mouse/human chimaeric B72.3 antibody. Prot. Engin. **1**; 499-505.

Wickens, M. (1990). How the messenger got it's tail. Addition of poly(A) in the nucleus. Trends Biochem. Sci. **15**; 277-281.

Widon, D. and Baldwin, R. L. (1980). Cation-induced Toroidal Condensation of DNA Studies with $Co^{3+}(NH_3)_6$. J. Mol. Biol. **144**; 431-453.

Wigler, M., Silverstein, S., Lee, L-S., Pellicer, A., Cheng, Y-C., Axel, R. (1977). Transfer of purified herpes virus thymidine kinase gene to cultured mouse cell. Cell **11**; 223-227.

Williams, G. V., Rohel, D. Z., Kuzio, J. and Faulkner, P. (1989). A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion-deletion mutants. J. Gen. Virol. **70**; 187-202.

Wolfe, J. A., Ludtke, J. J., Acsadi, G., Williams, P. and Jani, A. (1992). Long-term persistence of plasmid DNA and foreign gene expression in mouse muscle. Human Molecular Genetics **1**; 363-369.

Wynshaw-Boris, A., Short, J. M., Loose, D. S. and Hanson, R. W. (1986). Characterisation of the phosphoenolpyruvate carboxykinase (GTP) promoter-regulatory region. 1. Multiple hormone regulatory elements and the effects of enhancers. J. Biol. Chem. **261**; 9714-9720.

Yang, W. and Steitz, T. A. (1995). Crystal Structure of the Site-Specific Recombination γδ Resolvase Complexed with a 34 bp Cleavage Site. Cell **82**; 193-207.

Yarmolinsky, M. B. and Sternberg, N. (1988). The Bacteriophages Vol. 1, (Ed. Calender, R.) pp 291-438.

Yeung, C. Y., Ingolia, D. E., Roth, D. B., Shoemaker, C., Al-Ubaidi, M. R., Yen, J. Y., Ching, C., Bobonis, C., Kaufman, R. J. and Kellems, R. E. (1983). Identification of a functional murine adenosine deaminase cDNA clones by complementation in *E. coli*. J. Biol. Chem. **258**; 8338.

Yoshimura, F. K. and Chaffin, K. (1987). Different activities of viral enhancer elements before and after stable integration of transfected DNAs. Mol. Cell. Biol. **7**; 1296-1299.

Young, A. P. and Ringold, G. M. (1983). Mouse 3T6 cells that overproduce glutamine synthetase. J. Biol. Chem. **258**; 11260-11266.

Zachgo, J. and Gossler, A. (1993). Gene and enhancer trap screens in ES cell chimeras. Joyner, A. L. (Ed), In Gene Targeting. A Practical Approach, IRL Press, Oxford, pp 181-227.

Zentgraf, H., Frey, M., Schwinn, S., Tessmer, C., Willemann, B., Samstag, Y. and Velhagen, I. (1995). Detection of histidine-tagged fusion proteins by using a high-specific mouse monoclonal anti-histidine tag antibody. Nucl. Acids Res. **23**; 3347-3348.

Zhang, Y., Riesterer, C., Ayrall, A., Sablitzky, F., Littlewood, T. D. and Reth, M. (1996). Inducible site-directed recombination in mouse embryonic stem cells. Nucl. Acids Res. **24**; 543-548.

Zimarino, V. and Wu, C. (1987). Induction of sequence-specific binding of Drosophila heat shock activator protein without protein synthesis. Nature **327**; 727-730.