DEVELOPMENT OF VECTORS ALLOWING EFFICIENT HETEROLOGOUS-GENE EXPRESSION IN STABLE MYELOMA-CELL TRANSFECTANTS.

BY

KEVIN HUDSON.

A DISSERTATION SUBMITTED TO THE

UNIVERSITY OF LEICESTER

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

NOVEMBER 1988.

UMI Number: U010103

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI®
Dissertation Publishing

UMI U010103
Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest®

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITY LIBRARY LEICESTER

X751498245

CONTENTS.

Chapter 1: Introduction.

Chapter 2: Materials and Methods.

Chapter 3: Assessment of activity of expression elements in J558L.

Chapter 4.

High-level expression from transfected eucaryotic-genes in J558L.

Chapter 5: Expression interference in two-gene plasmids.

ABBREVIATIONS.

α	Alpha IgH-chain.
ATP	Adenosine 5'-triphosphate.
AMV	Avian myeloblastosis-virus.
Bal31	Bal31 exonuclease.
BiP	Immunoglobulin heavy-chain binding protein of ER.
Bis	Bis-acrylamide.
bp	Base pairs.
BPV	Bovine papillomavirus.
Brij 35	Polyoxyethylene-lauryl-ether.
BSA	Bovine serum albumin (fraction V).
Cμ	Constant region (of IgH(μ) gene and chain)
cad	N-phosphoacetyl-L-aspartate - resistance gene of *E.coli*.
CAT	Chloramphenicol acetyl transferase.
cDNA	Complementary DNA (i.e. made from a RNA template).
CHO	Chinese-hamster ovary cell-line.
CIP	Calf-intestinal alkaline phosphatase.
cPBS	complete PBS (PBS with 5mM $CaCl_2$ and $MgCl_2$).
CMV	Human cytomegalovirus strain-A169.
CMV-IE	CMV immediate-early region.
D	Diversity region (of IgH gene and chain).
(d)	Divergent two-gene expression vector.
Δ	Indicates the deletion of a restriction site or fragment.
ΔTK	Derivative of TK promoter lacking the 5' promoter-sequences.
dATP	2'-deoxyadenosine 5'-triphosphate.
dCTP	2'-deoxycytidine 5'-triphosphate.
ddATP	2',3'-dideoxyadenosine 5'-triphosphate.
ddCTP	2',3'-dideoxycytidine 5'-triphosphate.
ddGTP	2',3'-dideoxyguanosine 5'-triphosphate.
ddTTP	3'-deoxythymidine 5'-triphosphate.
dGTP	2'-deoxyguanosine 5'-triphosphate.
dTTP	Thymidine 5'-triphosphate
DEAE-	diethylaminoethyl-.
DMEM	Dulbecco's modified Eagle's-medium, containing 4.5g/l glucose and

	without sodium pyruvate.
DMSO	Dimethyl sulphoxide.
DNA	Deoxyribonucleic acid.
DNase	Deoxyribonuclease.
dNTP	Deoxynucleotide triphosphate.
DHFR	Dihydrofolate reductase.
DOC	Deoxycholic acid.
DTT	Dithiothreitol.
E-buffer	Tris-acetate-EDTA electrophoresis buffer.
EBV	Epstein-Barr Virus.
EDTA	Ethylenediaminetetraacetic acid (disodium salt).
EGTA	Ethyleneglycol-bis-(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid.
ELISA	Enzyme-linked immunoassay.
ER	Endoplasmic reticulum.
EtBr	Ethidium bromide.
FCS	Foetal calf serum.
γ	Gamma IgH-chain.
gpt	Xanthine-guanine phosphoribosyltransferase gene of *E. coli*.
H2A	Histone 2A.
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid (a buffer).
HGPRT	Hypoxanthine guanine phosphoribosyltransferase.
hgr	Hygromycin B gene of *E. coli*.
HS	Donor horse serum.
Ig	Immunoglobulin.
IgH	Immunoglobulin heavy-chain.
IgL	Immunoglobulin light-chain.
IMP	Inositol monophosphate.
IPTG	Isopropylthiogalactosidase
ISV40	DNA fragment containing small t-antigen intron and early polyadenylation-signal of SV40.
J	Joining region (of IgH gene and chain).
κ	Kappa IgL-chain.
Kb	Kilo base-pairs (for double-stranded nucleic-acid fragments)..
KD	Kilo daltons (for molecular weights).
Klenow	Large-fragment of DNA-polymerase I.

λ	Lambda IgL-chain.
LB	Luria-Bertani medium.
LTR	Long-terminal-repeat.
M13	Bacteriophage M13.
MET.	Methionine.
ML	Major late-region (of Adenovirus).
MMTV	Mouse mammary-tumour-virus.
Mops	3-(N-morpholino) propanesulphonic acid (a buffer).
mRNA	Messenger RNA.
MT	Metallothionein.
neo	Neomycin(G418)-resistance gene of *E.coli*.
NET	Buffer containing 150mM NaCl, 5mM EDTA and 50mM Tris.HCl (pH7.4).
NH_2	Amino group.
NP40	Nonidet P40.
OD	Optical density.
PAGE	Polyacrylamide-gel electrophoresis.
pBR	Sequences derived from pBR322.
PBS	Phoshate-buffered saline (without Ca^{2+} and Mg^{2+}).
PEG	Polyethylene glycol.
Pipes	Piperazine-N,N'-bis(2-ethanesulphonic acid) (a buffer).
Poly(A)	Polyadenylic acid sequence at 3' end of mRNA.
Poly(A)$^+$	Containing poly(A) tail.
Poly(A)$^-$	Not containing poly(A) tail.
RIA	Radioimmunoassay.
RIPA	NET containing 0.1% (w/v) SDS, 1.0% (w/v) Na DOC and 1% (v/v) NP40.
RF	Replicative form of M13 DNA (double-stranded).
RNA	Ribonucleic acid.
RNase	Pancreatic ribonuclease (Dnase free).
rpm	Revolutions per minute.
RSV	Rous-sarcoma virus.
snRNP	Small nuclear ribonucleoprotein.
S phase	DNA replication period of cell-cycle.
SDS	Sodium-dodecyl-sulphate.
SLR	Standard laboratory-reagent.
SSC	Solution containing 0.15M NaCl, 0.15M Na citrate (pH7.0).

SV40	Simian-virus-40.
(t)	Tandem two-gene expression vector.
TE	Buffer containing 10mM Tris.HCl, 1mM EDTA.
T4	*E. coli* infecting T4-bacteriophage.
TBE	Tris-borate-EDTA electrophoresis buffer.
TEMED	N,N,N',N'-tetramethylethylenediamine.
TK	Thymidine kinase (of Herpes-Simplex-Virus).
tRNA	Transfer RNA.
Tris	Tris(hydroxymethyl)-aminoethane (a buffer).
trp	*E. coli* tryptophan-gene.
μ	Mu IgH-chain.
μ_m	Terminus of μ-gene coding for a membrane integrating C-terminus of μ-chain.
μ_s	Terminus of μ-gene coding for a secretory μ-chain.
w/v	Percentage by weight (weight (g) in a 100ml volume of solution).
V_H	Variable (region of IgH gene and chain).
VA	Virus (adenovirus) associated.
v/v	Percentage by volume (ml) in a 100ml volume of solution.
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside.
XGPRT	Xanthine-guanine phosphoribosyltransferase.
XHMPA	Xanthine, hypoxanthine and mycophenolic acid (selective reagents for stable transfection).

ACKNOWLEDGEMENTS.

I would like to thank Dr. Tim Harrison for his advice and encouragement throughout the course of the project. I am also grateful for the help received from Professor W.J. Brammar and Dr. J. Sampson, who where involved in progress review meetings. Other members of the department provided valuable help and discussions, particularly Pete Jones, Glyn Price, Liz Allen, Sarah Munson and Jenny Ridge, and countless others who are to numerous too mention.

My thanks are due to my parents, without whose help and support during seven years of higher education, I could not have reached this stage of my career. Also, to friends and colleagues who provided light relief from the frustration and disappointment of experiments 'not working'. In particular to Maria, Miguel, Monica, Rick, Michelle, Paul and Fiona.

Finally I wish to acknowledge the support of the SERC for a Biotechnology-Directorate award, which allowed maintenance of some degree of solvency over the three years of study.

This thesis was completed in November, 1988 and submitted in accordance with the requirements of the University of Leicester for the degree of Doctor of Philosophy. Experimental work was carried out between October, 1985 and August, 1988.

PUBLICATION.

Some of the work described in this thesis has been published:

Hudson, K and Harrison, T.M. (1987) Stable myeloma cell transfection frequency variation dependent on different upstream expression elements. Biochem. Soc. Trans., 15, p1033-1034.

There is also a paper in preparation:

Hudson, K. and Harrison, T.M. Plasmid expression vectors for a mouse myeloma cell allowing high level expression of a recombinant protein.

CHAPTER 1: INTRODUCTION.

1.1. Heterologous expression-systems and their potential.

For several decades, the enzymic and signal-carrying functions of proteins have been exploited for use as therapeutic agents. The most suitable proteins are those which are secreted and carry out their function(s) extracellularly; for example proteins acting within or via the blood system, which can be administered to a patient intravenously. Best known examples of proteins of this type include insulin, somatotropin (growth hormone), urokinase, streptokinase and blood-clotting factors. However, the total number of proteins used has, until recently, been rather small considering that so many have potentially therapeutic properties. This was mainly due to their limited availability. They are usually too large and complex to be synthesised chemically, and many of them are present in the body only in trace amounts, making purification procedures both laborious and expensive.

In the late 1970s and the early 1980s, the advent of recombinant-DNA technology opened up the possibility of producing virtually unlimited quantities of any protein. The gene (or a cDNA) encoding a protein could be cloned, and then surrounded by active expression signals, such that on introduction into a suitable host-cell, the encoded protein could be synthesised at high levels. Unfortunately, the exploitation of heterologous expression-systems has not been as straightforward as first anticipated. Procaryotic-hosts have limitations with regard to expression of some eucaryotic proteins, and eucaryotic expression-systems are only a recent development. However, the Biotechnology market is now expanding. A list of proteins that have been produced in eucaryotic expression-hosts, which are either presently marketed or promising potential products, are listed in table 1.1.

Rather than producing the protein encoded by the native gene, many workers are attempting to improve on what nature has provided by synthesising 'second generation' proteins. The protein sequence can be modified through manipulation of the DNA coding-sequence, using techniques such as gene fusion and site-directed mutagenesis. For example, many second

Table 1.1. Proteins that can be expressed by recombinant methods in mammalian cells.

The table gives a representative list of proteins that have been expressed by recombinant methods in mammalian cells, and indicates their possible applications in the areas of human and animal health care. Some of these products are already commercially available, whereas others are only in the early stages of development. The list of such proteins is continuously expanding. The information in this table was collected from a series of articles; Spier, 1986; Ratafia, 1987; Ramabhadran, 1987; Van Brunt and Klausner, 1988.

Product	Disease /condition
Lymphokines Interleukins (2, and others) Interferons (α, β, γ) Tumour necrosis factor Colony stimulating factor	Viral infections / cancer
Thrombolytic agents Urokinase Tissue plasminogen activator	Thrombolysis
Vaccines Hepatitis B AIDS	Vaccination
Blood clotting factors (e.g VIII and IX) Growth hormone Protein C Insulin Erythropoietin	Supplementation of protein deficiency
Epidermal growth factor	Burns
Granulocyte stimulating factor	Wounds
Transfer factor	Multiple sclerosis
Orthaclone	Kidney transplant rejection
Chimeric antibodies	Site specific action of non-specific agents

generation tissue-type plasminogen activators have been produced, which potentially are clinically more efficient than the native protein (reviewed by Harris, 1987). Alternatively, chemical modifications can be made to the recombinant protein following its isolation from the expression host (Inada *et al*, 1986). For example, an acetylated derivative of streptokinase, which is marketed as 'Eminase', has increased plasma half-life compared with the native protein. Recombinant-DNA techniques are also being used to produce proteins which normally appear on the surface of infectious agents such as viruses and bacteria. When these proteins are administered to a patient, they may act as vaccines.

1.2. Requirements of an expression host.

The main requirements of a heterologous expression host are twofold. Firstly, it must synthesise a protein that maintains the biological and physicochemical features of the native protein (unless the aim is to purposefully change one of these features). The genetic code is almost universal and, therefore, the same DNA coding-sequence will generally encode a protein with an identical primary sequence of amino-acids in any host cell. However, the activity and physical properties of the protein may also depend on post-translational modifications (see section 1.10.), which may be cell-type specific. The fidelity of a large heterologous recombinant-protein to the native structure is in fact rather difficult to ascertain, particularly with respect to conformation. Therefore, there may be problems in convincing the government regulatory bodies that the proteins are safe for human therapeutic use. Consequently, as for any therapeutic agent, the suitability of a heterologous recombinant-protein to serve the desired role, in terms of both function and toxicology, must be evaluated by empirical determination in animal models and clinical trials. The second requirement of an expression host is that it expresses the protein at levels which make the system economically viable. For every protein, there is an optimum host-cell in which to produce it by recombinant means. In general, the host which is cheapest to grow, while expressing a recombinant protein which has properties similar to the native protein, is the most appropriate.

1.3. Expression hosts.

In the following sections, I will describe the range of expression hosts that are presently available for producing recombinant proteins, and the advantages and disadvantages of each system. They range from the simple procaryotic-hosts to the less easily manipulated mammalian hosts.

1.3.1. Bacteria.

At the time when recombinant-DNA techniques became available, *Escherichia coli* (*E. coli*) was the favoured host for expression of heterologous proteins. This was mainly because it was easily manipulated, and the rules governing gene expression were best understood in this organism. Consequently, expression vectors were relatively easy to construct. Some of the vectors now used in *E. coli* contain inducible expression-elements, which allow synthesis of the heterologous product to be switched on after the culture has become established (Renaut *et al.*, 1981). This may be beneficial, because synthesis of large amounts of a heterologous protein can be detrimental to cell growth. Other advantages of using *E. coli* include the low cost of growing the cells, and the high productivity of some heterologous proteins, which can account for up to 25% of total cell-protein (Caulcott and Rhodes, 1986). Several heterologous proteins produced in *E. coli* are available commercially. Examples are human insulin, human somatotropin and interferon α-2.

There has been a growing realisation, however, that *E. coli* is not the most suitable host for producing many eucaryotic proteins, particularly those which are large and complex. When most eucaryotic proteins are overexpressed in *E. coli*, they become denatured, and collect in structures known as inclusion bodies in which they become aggregated with nucleic-acids (Williams *et al.*, 1982). This presents not only the problem of purifying the protein, but also of solubilising and renaturing it. Renaturation can be achieved using agents such as SDS, guanidium HCl and urea, but it is often difficult to remove all traces of these agents. Also, if the protein is intended for internal human use, it is particularly important to remove all traces of endotoxins, which are components of the

E. coli cell-wall. The complex and laborious purification methods used for the isolation of eucaryotic proteins from *E. coli* have been reviewed by Marston (1986). Even if the heterologous protein is not denatured, it may still differ in conformation to the native protein. This is primarily due to the reducing environment of bacterial cells (Fahey *et al.*, 1977), which prevents the formation of disulphide bonds (see section 1.10.3.2.). In addition, proteolytic digestion of heterologous proteins produced in *E. coli* often occurs. This is because the heterologous proteins are recognised as foreign. This problem can be reduced by making a fusion protein of the heterologous protein with an endogenous protein (e.g., β-galactosidase-fusion-proteins), but there also needs to be a method for precisely excising the desired protein. Another potential problem is that all proteins in procaryotic cells are synthesised with a N-terminal, N-formylmethionine residue. The removal of the formyl group and/or the entire methionine residue from heterologous proteins may not occur, and its presence may affect the properties of the protein. However, it must be conceded that examples do exist where non-removal of the methionine has little or no adverse effect on eucaryotic proteins (Harkins *et al.*, 1981; Liang *et al.*, 1985).

It is desirable to secrete a heterologous protein, because this increases the likelihood of correct folding and also makes purification easier. However, *E. coli* does not usually secrete proteins. Even when a protein contains a signal-sequence (see section 1.10.1.), this usually directs the protein only to the periplasmic space, where it becomes concentrated. This may cause precipitation of the protein, and also make it more prone to protease digestion (Talmadgo and Gilbert, 1982; Chang *et al.*, 1988). Only a small number of proteins are secreted by *E. coli*, mainly enterotoxic proteins, colicins and haemolysin-A. Until recently, there were only a few examples of secretion of heterologous proteins from *E. coli* into the extracellular medium (e.g. Lord, 1985; Nagahari *et al.*, 1985), and such proteins usually accounted for less than 1% of total cell-synthesised protein. Recent progress on this front has involved the exploitation of factors which normally cause secretion from bacterial cells. Abrahmsén *et al.* (1986) and Uhlen *et al.* (1988) have shown that use of the Staphylococcal protein-A signal-sequence (and a short sequence normally downstream of it) will allow secretion of fused eucaryotic-polypeptides.

Another secretion system is based on the utilisation of lysis peptides, such as that encoded by the *kil* gene. This normally directs secretion of colicin E1, but has been used to allow secretion of other proteins which are normally delivered to the periplasm (Blanchin-Roland and Masson, 1988). Also, a C-terminal-region sequence of haemolysin-A, which is thought to be recognised by HlyB and HlyD proteins which mediate haemolysin-A secretion, has been incorporated into fusion proteins and shown to allow their secretion (Holland, 1988).

Other bacterial hosts are also being investigated as expression hosts. One prime candidate is *Bacillus subtilis*. It is non-pathogenic, does not contain endotoxins and can be manipulated by current recombinant-DNA techniques. Its major advantage over *E. coli* is its capacity to secrete a large number of heterologous proteins. However, the use of *B. subtilis* has so far been limited. This is due of the high concentration of intracellular and extracellular proteases in many strains, and also plasmid-instability in terms of both loss of the entire plasmid and genetic rearrangements (see Doi *et al.* (1986) for a review).

1.3.2. Yeasts.

Yeast is a unicellular, eucaryotic organism and is, therefore, regarded as an expression host whose complexity is intermediate between that of procaryotic and mammalian hosts. Yeasts are the only eucaryotic cell-type which can be manipulated biochemically and genetically (both classical and molecular) with the same ease as *E. coli*. Moreover, they may be a suitable expression-host for a wider range of eucaryotic-cell - produced proteins than bacterial cells (see Kingsman *et al.* (1987) for a review). For example, the recombinant vaccine, hepatitis-B-virus surface antigen, can be produced in yeasts (e.g. Cregg *et al.*, 1987), whereas problems associated with its expression in *E. coli* were intractable.

High-efficiency expression vectors are based on a yeast plasmid called the 2μM circle, and are stably maintained at 50 to 200 copies per cell (Broach, 1982). Transcription is driven by yeast expression-signals, such as the constitutive promoter from the phosphoglycerate-kinase gene (Tuite *et al.*, 1982), or the more inducible promoter from the GAL1 gene (St. John and Davis, 1981). Yeasts can secrete heterologous proteins which

contain their native signal-sequence, but secretion is often more efficient when a yeast signal-sequence is employed. They can also perform some post-translational modifications (see section 1.10.), such as N-acylation, fatty-acylation and phosphorylation. Hence, when simple proteins are required in large quantities, the ease and low cost of bulk yeast-fermentations may make them an attractive expression-host.

However, yeasts also have drawbacks as an expression host. The mode of glycosylation (see section 1.10.3.1.) deviates significantly from that of mammalian cells. Sambucetti *et al.* (1986) showed that the oligosaccharide side-chains tend to be of the high-mannose variety, which is often a precursor form of the more complex glycosylation-patterns found in mammalian proteins. Also, large and complex proteins have been difficult to secrete, to produce in high yields and to produce with full biological activity.

1.3.3. Fungal cells

Fungal expression-systems are still in the early stages of development, mainly because of the need to characterise expression-elements and to develop recombinant-DNA techniques for their manipulation. However, Gwynne *et al.* (1988) have developed an expression system using *Aspergillus nidulans*, in which sequences encoding proteins are linked to NH_2-terminal fungal secretion-signals, and expression is driven by the inducible promoter of the native alcohol-dehydrogenase-I gene. This has allowed the secretion of significant quantities of biologically-active proteins.

1.3.4. Insect cells.

Insect cell-lines, such as that derived from the ovarian tissue of *Spodoptera frugiperda* (fall armyworm, Lepidoptera, Noctuidae), in combination with Baculovirus expression-vectors, are promising expression hosts.

Transfer vectors, containing the highly active and regulated *Autographa californica* nuclear-polyhedrosis-virus (AcMNPV) polyhedrin promoter (Smith *et al.*, 1983a and 1983b) with a foreign gene inserted downstream, are cotransfected into the insect cells with wild-type

Baculovirus DNA, using the $CaPO_4$ method (see section 1.4.). Recombinant Baculoviruses are produced by homologous recombination, within the cell, and can be identified from the plaques they produce which have a distinctly different morphology from those of wild-type virus. The recombinant Baculovirus can then be used to infect an insect cell-culture, and the recombinant protein is produced 48-72 hours post-infection. Unfortunately, the culture is essentially lytic within 4-5 days, which necessitates the use of batch production. Even so, the recent development of low-cost, protein-free medium that supports insect-cell growth, virus replication and recombinant-protein expression, at levels equivalent to those in serum-containing medium (Inlow et al., 1987), and methods for large-scale insect-cell propagation in airlift fermentors (Inlow *et al.*, 1987), have advanced the commercial potential of this system. A wide range of recombinant proteins have been expressed at high levels and they are, in some cases, antigenically, immunologically and functionally, similar to their authentic counterparts (see Luckow and Summers (1988) for a review). However, a major limitation of insect cells may be the nature of protein glycosylation. Although studies are limited, it seems, as in yeasts, that high-mannose-type N-linked oligosaccharides initially added to the protein undergo limited modification, without the incorporation of sialic acid, galactose and fucose residues typical of many mammalian glycoproteins (Butters *et al.*, 1982; Kornfield and Kornfield, 1985).

1.3.5. Mammalian cells.

The above sections have described the use of non-mammalian-expression hosts, which are all capable of expressing some mammalian proteins. However, amongst the success stories are many less well publicised failures. For production of mammalian proteins, mammalian expression-hosts are the most likely to conserve the native biological-features and physiochemical properties. This is because they are the most likely to be able to reproduce the post-translational modifications (see section 1.10.) which occur on many mammalian proteins, and which may contribute to such properties. There was an initial reluctance to use mammalian cells, in culture, as expression hosts, because of the high cost of medium and the slow growth of the cells (Griffiths, 1986). However,

mammalian cells are only used when there is no effective alternative expression host and, therefore, these disadvantages are now accepted. Moreover, mammalian cell-culture technology is progressing rapidly. Many cells can now be fed on serum-free medium. Also, bioreactors are available in which cell concentrations of 10^6 cells/ml or higher are possible (Merten, 1987); cells in the tissues of the body are present at around 2 x 10^9 cells/cm^3.

For production of a biologically-active human therapeutic-protein, the human cell which normally produces the protein is likely to be the ideal host. However, primary cell-cultures are not suitable because the potential therapeutic protein is often expressed at extremely low levels and, besides, primary cell-cultures have a limited life *in vitro* and often lose the ability to produce the protein of interest (Eagle, 1965). Unfortunately, with the exception of a few tumour cells, it has been difficult to increase the lifespan of ('immortalise') human cell-lines in culture. They are not as readily immortalised by transfection of activated oncogenes or by use of chemicals as with murine cells, and they very rarely undergo spontaneous immortalisation. Also, in order to obtain high-level expression from cells, the development of a highly-efficient expression system is required. This requires methods for introducing the exogenous DNA into the cell, and also the availability of active expression-elements. Comparatively little work of this kind has been done on many human cell-types. Consequently, a suitable human-cell expression system is not always available for producing large amounts of a human therapeutic-protein.

In fact, the mammalian cell-lines used for the production of mammalian recombinant-proteins are, almost exclusively, derived from hamsters or mice. A series of expression-systems have been developed which can be used for producing a wide range of recombinant proteins in these host cells (see sections 1.4 - 1.8.). As discussed above, in such heterologous systems, the protein is much more likely to resemble the native protein than when produced in a lower-order expression host. Even so, not all the native features may be preserved, because many post-translational modifications show cell-type specificity (see for example, section 1.10.1.3.). Consequently, there is a need to develop a wider range of mammalian expression-hosts. This project has involved the development of

a mouse, myeloma cell-line. This is introduced in greater detail in sections 1.11. and 1.13.

1.3.6. Transgenic animals.

In the above sections, I have concentrated on heterologous-protein expression from cells grown in culture. However, it is also possible to produce heterologous proteins in transgenic animals. Genes can be introduced into the germ-line of several animals, usually by microinjection of DNA into fertilised eggs (Gordon *et al.*, 1980; Brinster *et al.*, 1981), and the eggs are then implanted into pseudopregnant females. The embryos may give rise to a transgenic animal in which all or many of the cells carry the foreign gene. A more recent alternative approach to making transgenic animals has utilised pluripotential stem-cells. For example, EK-cells which are a permanent cell-line of undifferentiated early-embryonic cells, derived by manipulation in culture of pre-implantation embryos, can be transfected with Retroviral vectors. A suitably transfected cell-line (e.g. containing a high copy-number of transfected sequences or homologously-recombined sequences) can be identified, and then implanted into a blastocyst. The cells recolonise efficiently in the embryo, and integrate uniformly to give widespread distribution in the resulting animals (Robertson *et al.*, 1986; Evans *et al.*, 1986; Evans, 1988). Through breeding of transgenic animals, progeny can be obtained in which all cells contain the transfected gene.

The exogenous genes usually show temporal and spatial specificity, in line with the normal specificity displayed by the expression elements flanking the coding sequence. Therefore, through use of appropriate expression-signals, it has been demonstrated that heterologous proteins may be specifically expressed from mammary glands and secreted into milk (Gordon *et al.*, 1987). Also, transgenic animals may offer the opportunity of introducing DNA into cell-types which are refractory to transfection (see section 1.4.), or where transfection methods have not been developed. For example, it appears that in order to produce the active γ-carboxylated form of the blood-clotting factor, factor IX, a liver derived host-cell-type may be the most efficient. Unfortunately, hepatocyte transfection systems are not yet well developed and, therefore, attempts

are being made to immortalise hepatocytes obtained from transgenic mice, which express recombinant factor-IX from the transfected gene (M. Courtney and J.-P. Lecoq, personal communication).

1.4. Introduction of exogenous DNA into mammalian cells.

Several different approaches may be used for introducing exogenous DNA into mammalian cells. This process is called transfection. This term is used rather than the term 'transformation', which is used to describe the introduction of exogenous DNA into a bacterial cell (see section 2.8.). This is to avoid confusion with the use of 'transformation' to describe the changes leading to uncontrolled cell-growth in mammalian cells. Since some cell-types are refractory to certain transfection procedures, the optimum method has to be determined empirically for each target-cell.

Many cells can be transfected using viruses. Foreign DNA can be introduced into a viral genome, and the resultant recombinant viral-genome can be packaged into a viral coat, for introduction into a host cell (see section 1.5.2.).

Transfection can also be obtained using naked DNA (DNA-mediated transfection). Adherent cells have traditionally been the most amenable to this type of transfection. Transfection of cells by addition of a calcium-phosphate coprecipitate of DNA (Graham and Van der Eb, 1973) is a simple and efficient method for fibroblastic and epithelial cell-lines. Cells may also be transfected by incubating in a DNA solution containing low concentrations of DEAE-dextran (Sompayrac and Danna, 1981), but this method was, until recently (Gopal, 1985), only used when transient expression was required. It seems that the method of electroporation, developed by Potter *et al.* (1984), will be the most universally applicable transfection method. In this method, a DNA solution containing a suspension of cells is subject to a high-voltage electric discharge. Holes or pores are apparently created in the plasma membrane (Kinositak *et al.*, 1977), allowing exogenous DNA to be taken up into the cell. Through optimisation of the electroporation conditions, stable-transfection frequencies of greater than 1% have been obtained (Chu *et al.*, 1987; Margolskee *et al.*, 1988).

Another transfection strategy is that of spheroplast fusion. In

this technique mammalian cells are fused, in the presence of polyethylene glycol (PEG), with spheroplasts prepared from bacterial cells harbouring the DNA of interest (Schaffner, 1980; Rassoulzadegan *et al.*, 1982). Oi *et al.* (1983) demonstrated that some lymphoid cell-lines, which had previously been resistant to several different transfection methods, could be transfected using spheroplast fusion. Therefore, in this study, I predominantly used spheroplast fusion as the transfection method for introducing DNA into myeloma cells.

1.5. Expression vectors for mammalian cells.

If DNA, containing a protein-coding sequence in isolation, is transfected into a cell, expression will only be obtained in a very small proportion of transfected cells. Therefore, before transfection into an expression host, a coding sequence is put into an expression vector. The expression vector contains transcription-control signals, which should drive high-level transcription of the protein-coding sequence. The vector also provides a means of identifying the cells which become stably transfected (see section 1.7.) or, alternatively, produces transfectants at such a high frequency that identification of stable transfectants among non-transfected cells is not essential. In addition, the sequences in the vector may allow propagation of the transfected gene within the host. In the following sections, I describe the different types of expression vectors which can be used in mammalian cells.

1.5.1. Plasmid expression-vectors.

A typical eucaryotic plasmid-vector contains three distinct regions. Firstly, it contains a bacterial sequence containing an origin-of-replication, which allows propagation of the plasmid in a bacterial host before transfection into animal cells. For this reason the plasmids are often called shuttle vectors. In addition, an antibiotic-resistance gene, also in the bacterial sequence, allows for selection of bacterial cells harbouring the plasmid. Using a bacterial host as a vehicle for the plasmid, the plasmid can be genetically manipulated, so as to improve its effectiveness when introduced into the mammalian host. The second set of

sequences in a plasmid vector are those from which expression may permit selection of the stably-transfected mammalian cells. Such sequences may also select for maintenance and/or allow amplification of the vector sequences (see sections 1.7. and 1.8.). Finally, there is the coding sequence for the protein of interest, which is flanked by regulatory sequences which should ensure its high-level transcription.

Plasmids vectors can be introduced into the host cell by either DNA-mediated transfection or spheroplast fusion, as described in section 1.4. The following sections describe several different classes of plasmid expression-vectors.

1.5.1.1. SV40-based vectors.

Some of the earliest expression vectors developed were based around the DNA tumour virus, Simian-virus-40 (SV40). Initially, fragments of the SV40 genome were replaced with equivalent sized fragments of foreign DNA. These recombinant genomes could be transfected into a host cell, following packaging into a viral coat (like those described in section 1.5.2.). In order for the recombinant genome to be replicated and packaged into virions, a helper SV40 was required, to genetically complement some of the functions lost from the recombinant genome during the replacement of viral sequences.

The above described system is now rarely used, and has been superseded by SV40-based recombinant-plasmid vectors. These are never packaged into virions and, therefore, there are no size constraints on the insert fragment. Also, they do not give lytic infection. They usually contain the SV40 origin-of-replication, and may also contain SV40 expression-signals, which function in a wide range of cell-types. The backbone of the plasmid is usually derived from the *E. coli* plasmid, pBR322.

In order to obtain high-level expression from SV40-based plasmids, following transfection into a host cell, a system allowing their autonomous replication, which increases their copy number, is desirable. This is the basis of the COS-cell system (Gluzman, 1981), in which the SV40-early protein, large T-antigen, is constitutively expressed. Large T-antigen supports the replication of transfected plasmids which carry the

SV40 origin-of-replication, to a copy number of 10^4 - 10^5 per cell. This system can only be used transiently, because the permissive cells are killed a few days following transfection. This problem was alleviated, however, by the development of temperature-sensitive COS cells, which express large T-antigen in a temperature-dependent manner. This allows for temperature-regulated plasmid replication and, hence, maintenance of stable transfectants with low copy-numbers of plasmid (Rio *et al.*, 1985).

A large number of expression plasmids have been constructed based around the plasmid, pSV_2 (Mulligan and Berg, 1980 and 1981; Southern and Berg, 1982). This plasmid contains an *E. coli* plasmid (pBR322) origin-of-replication, a SV40 origin-of-replication and a SV40 transcription-unit. This transcription unit contains the early promoter and enhancer (flanked by the late promoter) upstream and the early polyadenylation-signal (and the small t-antigen intron) downstream. A heterologous coding-sequence, from which expression is required, can be inserted between these upstream and downstream sequences. Such plasmids have been used in constructs designed to drive expression from the bacterial genes, *neo* and *gpt* (for example see pSV_2gpt, figure 5.1.), which act as dominant selectable-markers in animal cells (see section 1.7.). When transfected into non-permissive cells (cells which do not support plasmid replication), these markers allow selection for stably-transfected cells, in which the plasmid DNA is integrated into the host genome (see section 1.6.) and the selectable-marker gene is expressed. In theory, it is possible to obtain cell lines giving stable expression of any protein of interest by cotransfecting the encoding gene with a plasmid containing a selectable-marker gene (see sections 1.7. and 4.2.).

As plasmids such as pSV_2gpt are completely independent of any viral function (such as replication or packaging) for the generation of stably-transfected cells, and depend completely on chromosomal integration for their propagation, they have been called passively-replicated plasmid expression-vectors (Howard, 1983). In fact, the presence of SV40 sequences is irrelevant to the survival of such plasmids, and many passively-replicated plasmid vectors have been produced that do not contain SV40 sequences.

1.5.1.2. Bovine-papillomavirus - based vectors.

Sarver *et al.* (1981) developed shuttle-plasmid vectors, which contain a pBR322 derived backbone, and a large subgenomic fragment of Bovine-Papillomavirus-I (BPV) containing cloning sites for insertion of foreign DNA. When such plasmids are transfected into some cells (e.g. rodent fibroblasts), the BPV-replicon supports replication of the plasmid at fairly low levels. The plasmid can sometimes be stably maintained as an independently-replicating episome at 10-100 copies per cell, and drive expression of recombinant proteins at high levels (e.g. Pavlakis and Hamer, 1983). The BPV sequences responsible for maintenance of the the episome were identified by Lusky and Botchan (1984), and better defined expression vectors have since been constructed. The presence of the plasmids does not usually impair the host cell to any great extent, but transfectants of some cell-lines, such as mouse fibroblasts, C127 and NIH3T3, show altered morphology; they form foci, which consist of overgrown piled up cells within the background cell-monolayer, and this phenomenon can be exploited as a method for identifying transfectants. Transfectants, of other cell-types which will support episomal replication but which do not exhibit an altered morphology, can be identified by incorporation of a dominant selectable-marker gene into the BPV-based plasmid vector (reviewed by Campo, 1986).

Unfortunately, some workers have experienced problems with BPV-based plasmid vectors; Ashman and Davidson (1985) and Bostock and Allshire (1986) have shown that the vectors may undergo rearrangements and suffer high rates of mutation. This may be due to the extrachromosomal replication and/or the presence of foreign sequences. Also, the sequences often lose their episomal nature and become integrated into the genome (Macgregor and Burke, 1987). It remains to be investigated if BPV-based plasmid vectors are suitable for use in myeloma cells (Weidle and Buckel, 1987).

1.5.1.3. Epstein-Barr-virus - based vectors.

Plasmid vectors based on the Epstein-Barr-virus (EBV) replicon have been developed (Sugden *et al.*, 1985), and may have greater potential than BPV-based vectors. Human B-lymphocytes, that have previously been transformed with EBV and express the nuclear antigen, EBNA-1, can be transfected with plasmids carrying a *cis*-acting element of EBV called *oriP*

(Yates *et al.*, 1984). Such plasmids are stably maintained, episomally, at about 10 copies (1-60 range) per cell. A hygromycin-B - resistance gene (*hgr* of *E. coli*) on the plasmids acts as a dominant selectable-marker for stably-transfected cells. Sugden's plasmids were constructed primarily for studying viral replication. However, the plasmids have been modified by other workers. These plasmids contain further transcription units and can be used for heterologous-protein expression purposes. For example, Margolskee *et al.* (1988) have constructed a plasmid vector containing *oriP*, *hgr* and the EBNA-1 gene, and which also has a cloning site for a gene encoding a protein of interest. This does not require the use of pre-transformed cells. Such vectors can be used for cloning of mammalian genes by functional expression (see chapter 6).

1.5.2. Virus-based expression systems.

The plasmid vectors described in section 1.5.1. usually contain viral elements for propagation and/or function, but do not depend on packaging of virions at any stage in the expression system. This distinguishes them from the expression systems described below.

1.5.2.1. Retroviruses-based vectors.

The natural life cycles of Retroviruses (RNA viruses) have several features which make them attractive as expression vectors. The normal replication process involves the stable integration of a DNA copy of the genome into the host-cell genome, and this occurs in a precisely defined manner, via a long-terminal-repeat (LTR) sequence contained in the viral genome. Also, the infection process is normally non-toxic to the cell, in contrast to many other transfection procedures. In addition, Retroviruses can infect a very large number of cell types, and normally at a high frequency.

Only a small portion of the Retroviral genome is required for packaging of the genome into a protein coat and for viral DNA synthesis. The *gag*, *pol* and *env* genes of Retrovirus are unnecessary for these functions, and can be deleted and replaced with large (at least 10Kb) foreign-DNA fragments. This generates a recombinant genome in which the

foreign sequence is flanked by LTR sequences. In addition to sequences involved in the integration event, the LTR sequences also contain promoter and enhancer sequences to drive transcription. Other sequences adjacent to the LTR sequences, namely the ψ site and the OB site, allow encapsidation and DNA synthesis respectively. Infectious particles, containing the recombinant Retroviral-genome, can be produced by introduction of the recombinant genome into a starter culture of fibroblasts, which are superinfected with a helper virus. This is necessary because, the recombinant Retrovirus cannot replicate, due to the deletion described above; this function is provided in *trans* by the helper virus (Mann *et al.*, 1983). The recombinant virions generated can then be used to infect the expression host.

Following infection, and integration of the Retroviral vector into the host chromosome, the host cell can usually express the genes carried on the vector. Many vectors now contain a selectable-marker gene, such as *neo,* in order to identify stable transfectants (Hwany and Gilboa, 1984; Miller *et al.*, 1985), in addition to a cloning site for a gene of interest. The sequence encoding the protein of interest may be flanked by the Retroviral transcription-control elements of the LTR, which are active in a wide range of cell types. In some cases, however, non-viral expression signals may be more appropriate (Bernstein *et al*, 1985).

Despite the many advantages of Retroviral-based expression systems, there are often problems associated with their use. For example, there may be inadequate dual-expression from two-gene vectors for no immediately apparent reason (Brown and Scott, 1987). Also, the low copy-number of integrated Retroviral vectors (usually single copy) may explain why the expression level from the heterologous gene is often low (Cullen, 1987). In addition, some Retroviral vectors are genetically unstable and show a high spontaneous-mutation rate (Dougherty and Temin, 1986). Hence, the major application of Retroviral vectors seems to be for transfecting cells which are refractory to alternative transfection procedures and/or are only available in low numbers, such as haematopoietic stem-cells (Joyner *et al.*, 1983). For example, lineage-specific expression from an exogenous human - β-globin gene in murine recipients has been achieved, following reconstitution of bone marrow with Retrovirus-transduced stem cells (Dzierzak *et al.*, 1988; Mulligan, 1988). This is a promising step in

the development of gene therapy for disorders of haemoglobin production and other disorders of the blood system.

1.5.2.2. Vaccinia-virus - based vectors.

The pox virus, Vaccinia, has also been genetically manipulated to allow expression of heterologous proteins. The heterologous sequences are introduced into a region of the Vaccinia-virus genome that is non-essential for virus replication. Sequences as large as 25Kb can be inserted without any deleterious effects on virus growth (Smith and Moss, 1983). The viral genome is large and difficult to manipulate. Therefore, the foreign gene is normally introduced into the viral genome by targetted homologous-recombination. A plasmid, containing a sequence to allow homologous recombination, and also containing a foreign gene flanked on both sides by Vaccinia expression-regulatory sequences, is transfected (usually by $CaPO_4$ treatment) into a cell already infected with Vaccinia virus. Elegant selection-systems have been devised to identify the required recombinant-virus (e.g., Smith *et al.*, 1983). Vaccinia-virus - based systems are unique, because the vector DNA remains in the cytoplasm of the host cell rather than entering the nucleus. Hence, the foreign DNA is transcribed by Vaccinia-virus - encoded RNA-polymerase, which necessitates the use of a Vaccinia promoter and other Vaccinia expression-signals.

Vaccinia virus causes lysis of the cells and released viral particles can infect other cells. Thus, recombinants expressing foreign antigens can be employed as live vaccines for the immunoprophylaxis of infectious diseases of both human and veterinary importance (reviewed by Tataglia and Paoletti, 1988). The system is made more attractive for this purpose because of the success and expertise associated with the long-term use of Vaccinia virus for vaccination against smallpox. In fact, multivalent vaccines can be obtained, from Vaccinia-virus recombinants in which expression is driven from several heterologous genes (Perkas *et al.*, 1985). Despite their predominant use for vaccine production, Vaccinia recombinants can also be used to obtain large quantities of non-vaccine recombinant proteins. For example, the blood-clotting factor, factor VIII, has been expressed from a Vaccinia vector in hamster kidney cells (BHK 21) (Pavirani *et al.*, 1985).

1.6. Integration of exogenous DNA sequences into host-cell chromosomes.

For this study, it was decided to use passively-replicated plasmid expression-vectors. As discussed in section 1.5.1.1., such plasmids depend on integration into the host genome for their propagation. However, only scant details of the mechanism(s) by which exogenous DNA becomes integrated into the host genome, are understood. These are outlined below.

During the transfection procedure, a high proportion of cells may take up exogenous DNA, and often at high copy-number (as indicated by transient assays). The exogenous DNA is believed, in most cells, to become linked into large concatameric-units which have been termed concatamers or transgenomes (Scangos and Ruddle, 1981). The enzyme(s) involved in the production of concatamers has not been identified and, therefore, the mechanism by which the concatamers are produced is unclear. They might form by a mechanism of homologous recombination between the transfected plasmids (Folger *et al.*, 1982). However, Chia and Rigby (1981) showed that the formation of concatamers at least partially depends on plasmid replication. Another, less favoured, mechanism is one of end-to-end ligation of linearised transfected-sequences. The concatamers are unstable and are rapidly lost, but in a small fraction of the cells, a concatameric unit of exogenous DNA becomes integrated into the host genome (Perucho *et al.*, 1980). In some cells, however, only single copies of a transfected vector may be integrated into the genome. In most transfectants, the exogenous DNA is integrated at a single chromosomal-site (Robbins *et al.*, 1981; de Saint Vincent *et al.*, 1981; Folger *et al.*, 1982). The site of integration into the genome appears to be essentially random, and has been located near telomeres, centromeres, nucleolus organisers and in the euchromatin (Stark, 1984). Brenner *et al.* (1984) reported that the exogenous DNA integrated preferentially into repetitive DNA-sequences, but even here, there was no significant homology between the exogenous DNA and the integration site. However, integration of plasmid sequences by homologous recombination with complementary sequences in the chromosome can be identified. Smithies *et al.* (1985) showed that a transfected plasmid, containing a human β-globin locus, could homologously recombine with a chromosomal human - β-globin locus in a predictable fashion, to produce a sequence which was identified by a combination of selection and screening procedures. However, this

predictable event only occurred in approximately one in every thousand transfectants.

It has been shown that following integration of a vector into a host chromosome, the vector sequences can usually be stably maintained in a cell-line providing selection is maintained. If selection is removed, however, expression from the sequences may be lost (Gebara *et al.*, 1987 and see section 4.4.). This usually occurs at a low frequency, but seems to be sequence dependent; Retroviral sequences may be particularly unstable (see section 1.5.2.1.).

1.7. Selection for stable transfectants.

Usually (as alluded to in previous sections, and as is discussed in section 3.1.), only a small fraction of a population of transfected cells give rise to stable transfectants. Therefore, a selection system is required to identify stably-transfected cells amongst the background of non-transfected cells. In some cases, such as with BPV-based plasmid vectors (see section 1.5.1.2.), transfected cells can be identified by an altered morphology. However, this is not usually the case and, therefore, stable transfectants must be identified through expression from a transfected gene. This transfected gene encodes a protein which confers a dominant, selectable drug-resistance on the cells and/or complements an auxotrophic mutation.

Proteins encoded by some procaryotic genes can, in the presence of toxic drugs, act as dominant selectable-markers for transfection. These genes include *neo*, which confers resistance to G418, which is an analogue of the aminoglycoside antibiotic, neomycin (Southern and Berg, 1982), *gpt*, which confers resistance to mycophenolic acid (Mulligan and Berg, 1980, and see section 3.1.) and *hgr*, which confers resistance to hygromycin-B (Gritz and Davies, 1983).

Eucaryotic genes can also be used for dominant selection. This may be possible, even in cell lines which endogenously express the encoded protein. In such cases, a concentration of toxic drug is chosen which kills wild-type cells, but allows transfected cells to survive because of their higher levels of expression of the protein which confers resistance to the drug. The multifunctional enzyme, CAD, which contains aspartate-

transcarbamylase activity, confers resistance to N-phosphoacetyl-L-aspartate (de Saint Vincent *et al.*, 1981); the multidrug-resistance protein (a drug-efflux pump) confers resistance to vinblastine, colchicine, adriamycin and actinomycin D (Fojo *et al.*, 1985); adenosine deaminase (ADA) confers resistance to deoxycoformycin (Kaufman *et al*, 1986); glutamine synthetase (GS) confers resistance to methionine sulphoxime (Hayward *et al.*, 1986); metallothionein-I confers resistance to heavy metals such as cadmium (Hamer and Walling, 1982); and dihydrofolate reductase (DHFR) confers resistance to methotrexate (this is discussed further in section 1.8.). However, selection systems based on eucaryotic genes often work more efficiently when they are used to genetically complement an auxotrophic mutation in the host cell. Examples of this type of system include selection for dihydrofolate reductase (DHFR) activity in the DHFR$^-$ Chinese-Hamster ovary (CHO) cell-line, CHO-DUKX-B11 (Kaufman and Sharp, 1982); thymidine kinase (TK) in TK$^-$ mouse L-cells (Wigler *et al.*, 1979); hypoxanthine-guanine phosphoribosyltransferase (HGPRT) in HGPRT$^-$ cells (Szybalska and Szybalski, 1962); asparagine synthetase (AS) (actually from *E. coli*) in AS$^-$ cells (Holloman, 1975; Cartier *et al*, 1987); and ornithine decarboxylase (ODC) in ODC$^-$ CHO-cells (Chiang and McConlogue, 1987). .

Usually, there is no convenient selection system available for selecting cells stably transfected with a gene encoding a desired heterologous protein. Therefore, it is necessary to use a cotransfection approach. If an excess of a non-selectable gene of interest is mixed with a dominant selectable-marker gene prior to transfection, then these genes may become associated during the transfection process. Consequently, the non-selected gene is found to be present and expressed in some of the clones which have been selected for expression of the selectable-marker-gene product (Wigler *et al.*, 1979). In practice, simple cotransfection, in which the two genes are located on different plasmids, has not always been the most satisfactory method for cointegrating genes, especially into non-adherent cells (Gopal, 1985). Even when the non-selectable gene is present in large excess in the transfection (generally 10-20 fold), the level of expression, from the non-selectable gene in the resultant transfectants, is usually low. One reason for this is that even when cointegration of the selected and non-selected genes into the host chromosome occurs, usually only one or very few copies of the non-selected gene are present per genome

(Cullen, 1987). The cointegration frequency and average level of expression from a non-selectable gene can be increased by tighter linkage with the selectable gene, simply by cloning both genes into a single expression-vector.

1.8. Amplification of genomic sequences.

As alluded to above, there is a general correlation between the gene copy-number and the level of expression of the protein encoded by the gene. It must be emphasised, however, that this is only a loose correlation in the case of chromosomally-integrated genes; for example, Hanahan *et al.* (1980) could not detect any correlation between gene copy-number of integrated SV40-plasmid recombinants and the expression level of encoded T-antigen. The reason why the correlation is only weak, is that other factors such as chromosomal location also influence the level of expression (see section 1.9.1.3.).

The gene copy-number of genomic DNA sequences can, however, be increased by gene amplification (reviewed by Stark, 1984). Gene amplification refers to the production of additional copies of chromosomal sequence, which can be found as intrachromosomal or extrachromosomal DNA. This process allows cells to increase the expression of a gene product to a level necessary for survival, such as when exposed to a toxic drug. It has been applied particularly successfully to the DHFR selection-system (Kaufman and Sharp, 1982; see section 1.7.). The DHFR gene acts as a selectable marker for transfected cells by complementing a DHFR$^-$ auxotrophic-mutation in a CHO cell-line (CHO-DUKX-B11). Methotrexate (MTX) is an inhibitor of the DHFR enzyme, and when added in increasingly high concentrations to transfected cells, selects for cells which have undergone amplification of the DHFR gene(s). Flanking DNA sequences, including cointegrated non-selected genes, are co-amplified. Often, a DHFR gene copy-number greater than 100 per genome is produced (Kaufman *et al.*, 1985; McIvor *et al.*, 1985), and there is a report of 2,000 copies per cell (Crouse *et al.*, 1983). The reproducible high-level expression obtained from co-amplified non-selected genes makes the DHFR expression / selection / amplification system the most powerful system currently available for use in mammalian cells. One disadvantage of the system is that methotrexate is

very expensive. As methotrexate may have to be kept in the medium to select for maintenance of amplified sequences, this may result in large scale cultures being uneconomical to maintain. Also, the versatility of the system is limited due to the restricted availability of DHFR$^-$ cell-lines.

Several approaches have been used to make the DHFR selection / amplification system applicable to DHFR$^+$ cells and, therefore, a wider range of host cell-types. A mutant DHFR-cDNA can be used that encodes a DHFR protein with a reduced affinity for MTX. This is not inhibited significantly at concentrations of MTX which obliterate endogenous DHFR-activity and, therefore, kill non-transfected cells (Wigler *et al.*, 1980; Christman *et al.*, 1982; Simonsen *et al.*, 1983). Unfortunately, amplification of this mutant DHFR-gene is limited due to the high concentrations of MTX required. Another approach has been to use a very efficient DHFR expression-vector, which allows MTX to be added immediately following transfection, at a concentration that kills wild-type cells (Murray *et al.*, 1983). Also, the DHFR gene can be cotransfected with a dominant selectable-marker gene, which allows isolation of stable transfectants. Then the stable transfectants can be selected for DHFR-gene amplification by growth in MTX (Kim and Wold, 1985). Again, neither of these approaches has proven to be practical for convenient amplification of foreign DNA in all cell types.

Several other genes have also been used to select for vector amplification. For example, those encoding the following proteins: adenosine deaminase (Kaufman *et al.*, 1986); the multifunctional enzyme, CAD (Wahl *et al.*, 1984); glutamine synthetase (Hayward *et al.*, 1986); metallothionein-I (Hamer and Walling, 1982); thymidine kinase (Roberts and Axel, 1982); asparagine synthetase (Cartier *et al.*, 1987); multidrug-resistance protein (Fojo *et al.*, 1985; Gros *et al.*, 1986); and ornithine decarboxylase (Chiang and McConlogue, 1987) (see Bebbington and Hentschel, 1987 for a review).

1.9. Gene expression and its regulation.

Following introduction of an exogenous coding-sequence into a host genome, the level of expression from the coding sequence depends on the site of chromosomal integration, and also the vector sequences,

particularly those flanking the coding sequence. Some of the factors controlling expression have been identified, but very little is understood in detail about the mechanisms by which control is mediated. In order to ensure routine high-level expression from any heterologous gene, it will be essential to have a thorough understanding of gene expression control, so that expression systems can be designed rationally.

Some genes are expressed in a cell-specific manner, and this allows cells to develop and differentiate in order to perform specific functions in a multicellular organism. Often, the expression from genes which are expressed in a cell-specific manner is at a higher level, compared with the level of expression from housekeeping genes, which encode the functions necessary in all cells. In a cell which expresses a cell-specific gene, cell-specific factors are present to allow optimum expression from that gene. Therefore, it is desirable to use the cell-specific expression signals from such genes (or alternatively, powerful viral expression-signals), to allow, in the cell-type in which they are active, the potential for high-level expression from heterologous genes. However, this means that for each host-cell-type, one needs to identify the optimal expression-signals to incorporate into an expression vector. Such a strategy is the most rational presently available for ensuring high-level expression from a heterologous gene. Even so, each expression vector needs to be empirically tested to assess its efficiency, because there may be unforeseen influences on expression at various points in the expression pathway.

In the following sections, I wish to outline the multiple stages at which expression from any protein-coding gene, both endogenous and transfected, can be controlled in higher-eucaryotic cells. This includes transcriptional regulation, which has been extensively studied, and also post-transcriptional regulation which is less well characterised. Also, in section 1.12., a review of the control of IgH-chain expression is given, which serves as a specific example to illustrate how expression may be controlled at multiple points in the expression pathway.

1.9.1. Transcription.

Studies on a large variety of developmentally-regulated genes have indicated that transcriptional control is largely responsible for cell-specific differences in gene expression (Darnell, 1982). Hence, a high level of transcription is necessary in order to achieve optimum expression of the protein encoded by a gene. The following sections describe the factors that can influence the rate of transcription of a coding sequence.

1.9.1.1. Cis- and trans-acting transcription-regulatory elements.

Transcription of the majority of eucaryotic genes studied so far appears to be under negative control, which means they are inactive until transcription is switched-on. Once switched-on, the subsequent stages of control may be subject to positive or negative control. More documented examples of positive control have been identified, possibly because negative control is more difficult to study. The major interactions, determining the specificity and frequency of transcription initiation, are those between *cis*-acting DNA-sequence elements and *trans*-acting protein factors. For genes transcribed by RNA-polymerase II (protein-coding genes), the *cis*-acting sequences are subdivided into two classes, promoters or enhancers.

The promoter sequence usually controls the site of transcription initiation. It contains an AT-rich DNA-sequence motif called the TATA box, which is found 25-30 bases upstream of the transcription-initiation site (Breathnach and Chambon, 1981). Also, a consensus initiation-sequence, containing an A residue at which initiation occurs, and flanked by a pyrimidine residue on either side, has been proposed (Breathnach and Chambon, 1984). By definition, additional promoter sequence-elements, which also play an important regulatory role in the rate of transcription initiation, are confined within a region stretching no more than a few hundred base-pairs upstream of the transcription-initiation site. Examples include the 'CAAT' box and a GC-rich hexanucleotide (reviewed by McKnight and Tjian, 1986). However, in general there is a great deal of heterogeneity in the *cis*-acting regulatory-sequences upstream of the TATA box (Chambon *et al.*, 1984; Butcher and Trifonov, 1986).

Enhancers were originally identified in the genome of SV40 (Banerji *et al.*; Gruss *et al.*; Benoist and Chambon, all 1981), and are now

recognised as important elements controlling expression in a large number of cellular genes (Serfling *et al.*, 1985). Enhancers have been defined as sequences that dramatically potentiate (up to 1000 times) transcription from promoters; their effect is independent of orientation, position and, within limits, distance from the promoter. Comparisons of enhancer nucleotide-sequences have failed to reveal an absolutely maintained consensus, although many do possess a sequence similar to an eight nucleotide (GTGGAAAG) 'core' (Weiher *et al.*, 1983). Also, they often contain a repeated sequence-motif. For example, the SV40 enhancer contains a 72bp tandem-repeat. However, a repeat sequence does not seem to be a universal requirement of enhancers. Enhancers activate heterologous as well as homologous promoters and, therefore, the presence of an active enhancer in an expression vector is highly desirable for ensuring high-level transcription from heterologous genes.

In recent years, the distinction between promoter and enhancer elements has become less clearcut; they both seem to consist of multiple sequence-motifs, and some of these sequences may be interchangeable with no significant effect (Serfling *et al.*, 1985). In fact, many sequences, previously thought of as promoter elements, have been shown to function in an orientation-independent manner (Moye and Zalkin, 1987). Also, many of the motifs can be deleted without affecting expression, suggesting there is a considerable amount of redundant information in these *cis*-acting elements (Serfling *et al.*, 1985).

Experiments by Brent and Ptashne (1985) and Hope and Struhl (1986) showed that the yeast *trans*-acting proteins, GAL4 and GCN4, contain domains which possess DNA-binding or transcription-activating properties. These domains are remarkably independent of each other. Deletion of the activator domains leaves DNA binding activities intact, and when the activator domains are fused to the bacterial DNA-binding protein, LexA, the fusion-proteins enhance transcription from promoter-reporter constructions downstream of LexA binding-sites. It appears that most *trans*-acting transcription-regulating factors have this modular structure. The DNA-binding domain is a well-defined structural motif, and usually contains one of two domain structures, known as helix-turn-helix or zinc-finger domains (reviewed by Evans and Hollenberg, 1988). In contrast, the transcription activating (or repressing) domain appears to have a less well-defined

structure. In fact, mutational analysis of this domain suggests the only requirement is a sufficient excess of acidic residues, and this has led to them being called 'acid blobs' or 'negative noodles' (Sigler, 1988).

It appears that *trans*-acting factors may be cell-type specific, or only available for binding in certain cells, and it is this property which leads to cell-specificity of expression. In fact, some of the factors binding to *cis*-acting elements seem to be of a general type as well as cell specific; for example, the well characterised factor, Sp1 (Dynan and Tjian, 1985, and for a review, see Kadonagen *et al.*, 1986). A cassette-shuffling model (Voss *et al.*, 1986) proposes that binding of a combination of specific and non-specific factors by the enhancer, from a cellular pool, determines cell-specificity of expression.

The action of enhancers is further complicated by the observation that the binding of some *trans*-acting proteins serves to negatively regulate transcription. For example, reduction of early strand RNA synthesis of SV40 is inhibited by SV40 large T-antigen (Gaub *et al.*, 1985). Also, E1A proteins of Adenovirus 2 have the ability to both positively and negatively-regulate transcription, by their interaction with different promoters (Jones, 1986). Examples of endogenous negative-regulation on cellular genes have also been identified, and the cell-specific nature of the effect strongly implies that this is also mediated by *trans*-acting factors (Sassone-Corsi and Borrelli, 1986, and see section 1.12.2.).

The evidence for involvement of *trans*-acting factors in controlling transcription is often indirect, and their analysis has been limited because they have been difficult to purify. However, rapid progress is now being made in this area, as illustrated by a recent compilation of transcription-regulating proteins (Wingender, 1988). Cloning of the genes encoding the factors is expected to rapidly advance understanding in the near future.

1.9.1.2. Mechanism of enhancer action.

It is believed that a mechanism, more complex than simply binding of *trans*-acting factors to *cis*-acting regulatory sequences, is necessary for the transcription-initiation rate to be influenced. The numerous models proposed to account for the biological properties of

enhancers have been summarised by Corney *et al.* (1986) and Jeang and Khoury (1988). Many of the mechanisms are by no means mutually exclusive. Indeed, a combination of several mechanisms may operate, or different enhancers may operate by different mechanisms, because existing experimental evidence does not verify a universal mechanism.

The chromatin structure associated with enhancers is often nucleosome-free and hypersensitive to nuclease digestion. These regions have been called DNaseI-hypersensitive sites (Yaniv and Cereghini, 1986). For example, the 72bp tandem-repeat sequence of the SV40 enhancer is DNaseI-hypersensitive (Jongstra *et al.*, 1984), and this correlates with electromicrographic findings of a nucleosome-free region surrounding this enhancer (Saragotti *et al.*, 1980).

In vitro studies on the SV40 enhancer show it can assume a Z-DNA configuration (Nordheim and Rich, 1983), instead of the more normal B-DNA configuration. This suggests that the enhancer may act by organising the chromatin into a transcriptionally-active conformation (the chromatin-structure model). A related model (the site-specific gyrase model) suggests that enhancers serve to modulate the superhelicity of linked DNA, by providing sites for the action of topoisomerases. This is supported by *in vivo* studies, which show that topoisomerase-II cleavage sites are located adjacent to the SV40 enhancer (Yang *et al.*, 1985). In addition, topoisomerase-II inhibitors can inhibit RNA-polymerase-II - dependent transcription (Ryoji and Worcel, 1984). However, evidence for the above two topological models is incomplete. For example, it is contradicted by the finding that when the SV40 enhancer is conformationally uncoupled but not physically separated from a reporter gene, the transcription rate appears to be unaffected (Plon and Wang, 1986).

Other models suggest that the site of enhancer binding acts as an entry-site or 'open window' for RNA-polymerase II. It is proposed that the enzyme then finds the transcription-initiation site by scanning along the DNA sequence (the scanning model). This is supported by the observation that the positive effect of some enhancers (e.g. SV40) on distal promoters is reduced by the presence of a proximally-interposed promoter, or by a second promoter located on the opposite side of the enhancer (Kadesh and Berg, 1986). Also, when sequences between a SV40 enhancer and a downstream reporter-gene are chemically cross-linked, which might be expected to

interfere with scanning by RNA-polymerase II, the stimulatory effect of the enhancer is considerably reduced (Courey *et al.*, 1986). Again, this model does not appear to be a universal mechanism, because the effect of several enhancers on distant promoters is not inhibited by interposed promoters. For example, the Ig-κ enhancer can equally activate two κ promoters, despite their physical arrangement, in which one promoter is interposed between the enhancer and the second promoter (Atchison and Perry, 1986).

The enhancer effect is usually detected following gene transfer into an appropriate cell-line (i.e. *in vivo*). However, the effect has been difficult to detect in *in vitro* transcription-assays, using nuclear extracts from the same cell-type. This suggests that a higher-order nuclear structure, such as attachment to a nuclear compartment, is required (the nuclear-address model). One possibility is that the enhancer sequence coincides with specific attachment-points of DNA to the chromosomal scaffold or matrix (Mirkovitch *et al.*, 1984; Reeves, 1984; Jackson, 1986).

With available evidence, the most favoured hypothesis is one which proposes that transcription is stimulated by interactions between proteins which are bound to the enhancer or the promoter. It is suggested that this allows formation of a transcriptional complex for RNA-polymerase II to enter, with looping-out of intervening DNA (the looping model) (Ptashne, 1986). This model is supported by data which suggests a requirement for strategic spacing of SV40-enhancer motifs (Takahashi *et al.*, 1986); in turn, this suggests binding factors must be stereospecifically aligned to form a large interacting-complex which is located primarily on one side of the DNA-helix. In support of the looping-model, there is precedence for DNA-bending or DNA-looping in molecular interactions between proteins and DNA; for example, site-specific recombination and initiation of DNA replication in procaryotes (Echols, 1986). More recent hypotheses based on the same principles of protein-protein interactions leading to transcription control, have been made by Guarente (1988) and Sigler (1988).

1.9.1.3. Influence of chromosomal location on transcriptional activity.

In addition to the nature of the immediate flanking-sequences, the level of loading of RNA-polymerase onto a gene is also likely to be

greatly influenced by the chromosomal location. This, and also the gene copy-number, most likely account for the large variation in expression levels usually observed between transfectants, which have been transfected with the same expression vector. For example, Jaenisch *et al.* (1981) showed that the site of integration was often the determining factor for expression from Retroviral expression-vectors.

Chromosomal effects may be explained by large-scale differences between regions of chromatin; euchromatin is open and considered to contain transcriptionally-active sequences, whereas heterochromatin, which is condensed, is generally regarded as inactive. The phenomenon of X-chromosome inactivation clearly demonstrates the relationship between genetic inactivity and highly condensed chromatin. The most striking visible example of differential chromatin-packing is the banded appearance of dipteran polytene-chromosomes, where highly-condensed chromatin bands alternate with dispersed chromatin-bands. On gene activation, individual bands or groups of bands disperse to form puffs. The relationship between cytological structure and underlying genetic organisation was endorsed in a recent study of 315Kb of continuous sequence of a *Drosophila* chromosome. This sequence was cloned (by chromosome walking), which allowed probes to be made. These were then used to measure the transcription rates from the genes along the fragment. In salivary glands, the variation in transcription rates along the fragment approximately paralleled the distribution of bands over the same fragment in polytene chromosomes (Bossy *et al.*, 1984).

More localised differences in chromatin structure may also affect gene expression. For example, in the vast majority of cases studied in eucaryotes, there appears to be an inverse correlation between the level of DNA methylation, in defined areas around genes, and transcriptional expression (Doerfler, 1983). The preferred substrate for the methylase enzyme, which methylates cytosine-bases to produce the 5-methylcytosine derivative, is newly replicated hemi-methylated DNA. Hence, the methylation pattern is maintained through cell-divisions. Many housekeeping genes are flanked by sequences which contain a high frequency of CpG dinucleotides. Such sequences are called *Hpa* II tiny-fragments (HTF), because they were originally detected by digestion with the methyl-sensitive restriction enzyme, *Hpa* II. If the cytosine bases of these dinucleotides are densely

methylated, gene transcription may be repressed (Bird, 1986). Hence, integration of a transfected gene into such a region might result in low levels of transcription.

Chromosomal DNA in eucaryotes is bound to a group of small basic-proteins called histones. There are five major types of histones; H1, H2A, H2B, H3 and H4. However, each type of histone can exist in a variety of forms. Hence, variation in the histone isoforms within localised regions of chromatin might have a significant effect on chromatin structure and, therefore, affect the transcription rate of nearby genes. There are subsets of each histone type which are encoded by different genes. The extent of the differences ranges from single amino-acid substitutions in the highly-conserved H4, to the extensive deletion, insertion and substitution variants of H2A (Wu *et al.*, 1986). H1 is the most variable of all histones, and subtype H1° might be an especially relevant variant because it is distributed in a non-random manner within chromatin with respect to active, inactive and inducible gene-sequences (Mendelson *et al.*, 1986). In addition, histones can be extensively modified post-translationally, by methylation, acetylation, phosphorylation, poly(ADP)-ribosylation and ubiquitination. Most of these have been proposed to alter chromatin structure by disrupting DNA-protein or inter-fibre interactions, so 'loosening' the structure for template function (see Blanchard (1987) for a review).

A 'chromosomal' effect can also be caused by integration adjacent to an active, endogenous transcription-control element. Integration adjacent to an active enhancer can increase the level of expression (Sorge *et al.*, 1984; Taketo *et al.*, 1985), whereas integration adjacent to a negative regulatory-element, or 'silencer', can inhibit expression (Laimins *et al.*, 1984). Recently, Grosveld *et al.* (1987) demonstrated, initially with mouse β-globin mini-genes, and then with heterologous genes, that sequences can be integrated into the genome of transgenic mice and expressed specifically in erythroid cells at high levels, without the deleterious position-effects previously found to affect integrated genes. This is dependent on the genes being flanked, in the expression vectors, by sequences which contain DNAse-I-superhypersensitive sites, and which are normally positioned 50Kb upstream and 20Kb downstream of the β-globin gene, in the flanking regions of the globin locus. It is proposed

that these sequences activate the entire globin locus in erythroid cells, and more localised control sequences determine the temporal specificity of expression of the different globin-genes within the locus (Grosveld, 1988). Sequences with similar properties have also been identified in the thymus-cell-specific gene encoding the human T-lymphocyte marker, CD2 (T11) (Kioussis, 1988). It is likely that many other genes, whose expression is tightly controlled and cell-specific, are flanked by such sequences.

In conclusion, position effects are only likely to affect the average level of gene expression from transfected sequences, and should not preclude the isolation of high-level expressing clones. Providing the site of integration is random, screening should allow identification of clones in which the site of integration is favourable for high-level transcription.

1.9.1.4. Premature transcription-termination.

For most mammalian protein-coding genes studied, it seems that following transcription initiation by RNA-polymerase II, the polymerase produces a complete primary transcript by progressing to the 3'-end of the gene. Therefore, the amount of full-length transcripts produced is largely determined by the rate of transcription initiation by RNA-polymerase II. Even so, some mechanisms which mediate premature transcription-termination have been implicated in controlling expression in eucaryotic cells. For example, in animal viruses, SV40 and Adenovirus, attenuation may regulate gene-expression (reviewed by Aloni and Hay, 1985), and premature termination also seems to play an important regulatory role in the expression of the cellular proto-oncogene, c-*myc* (Bently and Groudine, 1986). Although such mechanisms do not appear to be as widespread as in procaryotes, it is obviously important to avoid the introduction of sequences into heterologous transcription-units which cause premature transcription-termination.

1.9.2. 5'- and 3'-processing of the transcription product.

Soon after initiation of transcription by RNA-polymerase II, the RNA is usually capped at the 5'-end (see Fraser *et al.*, 1979). This involves modification of the initial nucleotide of the primary transcript by addition of a 7-methylguanosine residue, which is linked by an unusual 5'-to-5' triphosphate bridge. Also, a methyl group is added to the 2'-OH group on the ribose sugar of the initial nucleotide of the transcript, and the second (N+1) nucleotide of the transcript is sometimes methylated. Capping is (except in Picornaviruses) crucial for activity of mature mRNA, because it is specifically recognised by ribosomes as an initiation signal for protein synthesis (Shatkin, 1976). In addition, the cap structure may contribute to the stability of mRNAs by protecting them from 5'-exonucleolytic degradation (Furuichi *et al.*, 1977; Shimotohno *et al.*, 1977). It has also been observed that RNA splicing (see section 1.9.3.) occurs more efficiently when the RNA transcript is capped, suggesting that cap recognition is an important step in the formation of a specific ribonucleoprotein (RNP) complex required for splicing (Krainer *et al.*, 1984; Konarska *et al.*, 1984; Edery and Sonenberg, 1985). However, it is not known if the addition of the cap is a regulated process which plays a significant role in controlling expression.

In higher eucaryotes, the 3'-ends of mRNAs are not generated by termination of transcription, as is the case in procaryotes. RNA-polymerase-II transcription generally transcription downstream of the functional mature 3'-end of the mRNA, and so the end is generated by post-transcriptional processing. This post-transcriptional processing (see Brawerman, 1981) involves an endonucleolytic cleavage at a specific site in the primary transcript, followed by immediate addition of a polyadenylate (poly(A)) tail, of relatively homogeneous-length (200-300 nucleotides), to the free 3'-OH end of the cleaved transcript.

Bone fide transcription termination by RNA-polymerase II has been shown in only a handful of systems (see Friedman and Imperiale (1987) for a review). Even in these cases, an exact termination site has not been identified, probably because termination occurs at multiple sites in a stretch of DNA extending over hundreds or even thousands of nucleotides. The sequence motifs containing the termination signal seem to be separated by non-essential sequences, as shown by Johnson *et al.* (1986) for the sea-urchin H2A gene. Termination sites seem to be dispensable, because

expression vectors that do not contain recognised terminators can still give efficient expression, and in such cases termination probably occurs when the polymerase falls-off the coding strand non-specifically.

For most mRNAs (histone mRNA being an exception), the addition of a poly(A) tail to the 3'-end appears to be essential, in order to give efficient expression of the protein encoded by the gene. Without polyadenylation, a nascent-RNA chain may be unstable (see section 1.9.5.) with a half-life of less than 10 minutes (Citron *et al.*, 1984). Sequences which contain the signals defining the site of cleavage and polyadenylation have been identified in a large number of genes. A highly conserved sequence (AAUAAA) is found just upstream of most eucaryotic mRNA 3'-ends (Proudfoot and Brownlee, 1976). More recent mutational analysis has revealed that sequences downstream of the mature 3'-end are also required for efficient cleavage of the primary transcript (Conway and Wickens, 1985). A consensus sequence has been suggested by McLauchlan *et al.* (1985) for these downstream sequences, but many downstream regions do not conform to the consensus and so it appears there are different classes of poly(A) signal.

It is likely that *trans*-acting factors interact with the *cis*-acting sequence elements described above to promote the events of 3'-processing, perhaps akin to the way in which transcription is initiated. For example, a specific small-nuclear ribonucleoprotein (snRNP) has been shown to associate with the AAUAAA sequence of RNA transcripts (Hashimito and Steitz, 1986). Most importantly, 3'-processing may show cell-type specificity. Thus, specific sequences may be required to allow efficient production of heterologous poly(A)$^+$-RNA and, consequently, increase the potential for high-level protein expression. However, the importance of using specific 3'-processing elements is an unresolved issue. For interferon genes, the 3'-processing elements have been shown to dramatically influence expression (van Heuvel *et al.*, 1986), but other studies have found little difference in expression from a gene using different 3'-processing signals (e.g. Weidle and Buckel, 1987).

It is possible that expression might be increased by the presence of a transcription-terminator downstream of the poly(A) signals. This could prevent transcriptional interference between adjacent transcripts (see chapter 5). In addition, there is evidence which suggests the processes of

termination and 3'-processing can be coupled. For example, Falck-Pedersen *et al.* (1985) demonstrated that the mouse β-globin gene termination-process requires not only the region of 3'-flanking sequences where termination is observed, but also sequences including the major poly(A)-signal. Also, Whitelaw and Proudfoot (1986) showed that a case of α-thalassaemia is due to a point mutation in the poly(A) signal, which results in the cleavage and polyadenylation reactions being very inefficient. This deletion also results in transcription proceeding far downstream from its normal end-point. In addition, Lanoix and Acheson (1988) showed that when a fragment, containing the rabbit β-globin polyadenylation-signal, is substituted for the poly(A) signal of Polyomavirus late-region, it ensures efficient polyadenylation of late-region transcripts, and simultaneously increases the efficiency of transcription termination. The examples above demonstrate that a functional poly(A)-signal can be part of a termination signal. On the other foot, studies of SV40 transcription in *Xenopus* oocytes indicate that polyadenylation is inefficient when termination efficiency is reduced (Miller *et al.*, 1982). Another possible reason for including an efficient terminator is that when termination is inefficient, transcription of the gene might be reduced due to retarded cycling of a limited amount of RNA-polymerase II.

In addition to the consideration of which 5'- and 3'-flanking sequences lead to the most efficient processing events, it must also be remembered that these sequences encode part of the noncoding regions present in the mature mRNA. These flanking sequences can influence the stability (e.g. Rabbits *et al.*, 1985; Morris *et al.*, 1986) and/or translation efficiency (e.g. Pelletier and Sonenberg, 1985) of the mRNA. These topics are discussed in sections 1.9.5. and 1.9.6. respectively.

1.9.3. RNA-splicing pathways.

The majority of eucaryotic genes contain intervening sequences which must be spliced out of the primary RNA-transcript, to produce the mature mRNA which is colinear with the protein sequence (Gilbert, 1978). The components of the splicing reactions have been partially characterised. The reactions are believed to occur in RNP complexes called splicesomes (Grabowski *et al.*, 1985; Perkins *et al.*, 1986), which contain several

small-nuclear RNPs (snRNPs). The snRNPs are complexes of RNA (five species of RNA have been implicated in splicing; U1, U2, U4, U5 and U6) and numerous proteins. Introns of higher eucaryotic pre-mRNA are bordered by consensus sequences (C/A)AG'GU(G/A)AGU at the 5'-end and $(C/U)_{11}$-NCAG' at the 3'-end (Mount *et al.*, 1982; Brown, 1986). Cleavage takes place at the 5'-end of the GU and the 3'-end of the AG sequences underlined. *In vitro* studies have shown that splicing proceeds in two stages. Firstly, cleavage occurs at the 5'-splice site. This may be mediated by a small, nuclear RNA (U1) which shows striking complementarity to the 5'-splice-site consensus (Lerner *et al.*, 1980; Garabowski *et al.*, 1984). The 5'-terminus of the intron then becomes bound to the 2'-OH group of an adenine-base located in the intron sequence, to form a lariat or branch-like structure (Padgett *et al.*, 1985). The second step of the splicing reaction involves cleavage at the downstream splice-site, and then ligation of the 3'-end of the upstream exon to the 5'-end of the downstream exon.

Despite the partial unravelling of the splicing mechanism, and the identification of the components involved, little is known about the control of the process. There is a lack of knowledge, particularly with respect to how passage through a splicing pathway may affect the rate and efficiency of mature mRNA production. Even so, this pathway clearly has the potential to be regulated. In the past couple of years, three cases have been identified in which expression of *Drosophila* proteins is turned on and off by controlling splicing-events necessary to produce the corresponding mRNA (reviewed by Bingham *et al.*, 1988). The importance of introns for efficient gene-expression in mammalian cells is an unresolved issue. Early work with recombinant SV40 showed, convincingly, that efficient formation of 16S mRNA requires the presence of an intron (Gruss and Rhoury, 1980), suggesting that passage through a splicing pathway might be a general requirement. However, there is now increasing evidence that introns are not always essential. For example, a recent carefully-controlled study came to the conclusion that introns are inconsequential to expression levels from a thymidine kinase (TK) gene (Gross *et al.*, 1987).

1.9.4. RNA transport.

RNA transcription takes place in the nucleus. For the mRNA to be translated, it must be transported to the cytoplasm. Current evidence suggests that RNA is never in a freely diffusible form; it seems likely that sequential attachment and detachment processes, at specific binding sites throughout a 'RNA transport' pathway, make the transport of RNA a regulated process (Schröder *et al.*, 1987). Consequently, expression levels might be influenced at the level of RNA transport. However, the phenomenon of RNA transport has not been studied sufficiently, to identify specific cases in which expression levels are primarily controlled at the level of RNA transport.

Essentially all heterogeneous-nuclear (hn) RNA (RNA prior to maturation into mRNA) is associated with a skeletal nuclear-substructure called the nuclear matrix (Faiferman and Pogo, 1975; van Eeklen and van Venrooij, 1981). This is where transcription itself takes place (Jackson and Cook, 1985). The mRNA must usually be matured before it can be transported to the cytoplasm. This suggests there must be some selection system to release mature mRNA from the nuclear matrix. Schröder *et al.* (1987) have shown that the release of mature hen-ovalbumin mRNA from the nuclear matrix can be caused by the presence of ATP and its non-hydrolysable analogues, and this can be inhibited by topoisomerase inhibitors. This suggests that changes in DNA secondary-structure may be associated with detachment of mRNA from the nuclear matrix. The second stage of RNA transport is the passage of the mature mRNA across the nuclear envelope into the cytoplasm. This is thought to be mediated by a nuclear-envelope-bound NTPase (Schröder *et al.*, 1986). This enzyme is stimulated by polyadenylic acid, and this has led to the suggestion that binding to the poly(A) tail of mRNA is involved in the translocation process. On entry into the cytoplasm, the mRNA binds to cytoplasmic skeletal elements. It has been suggested that microtubules and actin filaments (Schröder *et al.*, 1982) and/or intermediate filaments (Bachmann *et al.*, 1986) are involved in mRNA metabolism and transport.

1.9.5. mRNA stability.

The expression level from any gene is dependent on the steady-state concentration of the corresponding mRNA in the cytoplasm. As

discussed in section 1.9.1., the mRNA concentration is largely determined by the rate at which the DNA coding-sequence is transcribed. However, mRNA stability is also likely to influence considerably the steady-state level of expression. Consequently, a high transcription-rate will be irrelevant (with regard to obtaining constitutive high-level protein expression) if the mRNA is rapidly degraded once it enters the cytoplasm. The half-life of mRNA in mammalian cells varies widely; β-globin mRNA seems to be extremely stable, and many mRNAs have half-lives of hours or even days (Volloch and Housman, 1981). In contrast, some mRNAs are very unstable, such as those from the transiently expressed c-*fos* and c-*myc* genes, which have half-lives of around 30 minutes.

Raghow (1987) emphasised that there are three different influences on mRNA stability. Firstly, endogenous and exogenous stimuli can trigger changes in mRNA stability. For example, the level of histone mRNA fluctuates during the cell-cycle. The histone mRNA concentration increases during DNA synthesis (S-phase), and this is partially due to a temporal increase in stability (Graves *et al.*, 1987). Secondly, there may be a relationship between protein synthesis and mRNA turnover (discussed by Brawerman, 1987). For example, Graves *et al.* (1987) reported that histone mRNA is more susceptible to degradation when it is being translated. Also, β-tubulin mRNA seems to be made susceptible to degradation by a mechanism which recognises the first four amino-acids of the nascent β-tubulin chain, as they emerge from the ribosome (Yen *et al.*, 1988). Thirdly, and the most relevant feature of mRNA stability to the work described in this thesis, is that each mRNA has an intrinsic susceptibility to mRNA degradation. Degradation is thought to be mediated by specific factors which recognise unique sites on mRNA chains (Brawerman, 1988). It has been suggested that a rate-limiting endonucleolytic cut, triggered by such an interaction, could be followed by rapid destruction of the mRNA by exonucleases. A 3'→5' exonuclease, identified by Ross *et al.* (1987), is probably at least partly responsible for this process. However, as discussed in section 1.9.2., the 5'-cap structure seems to provide protection against 5'→3' exonucleolytic activity. An AU-rich sequence, promoting mRNA-decay, has been identified in the 3'-noncoding region of a human lymphokine RNA (GM-CSF) (Shaw and Kamen, 1986), and destabilising sequences have also been found in the 3'-noncoding regions of the c-*fos* and c-*myc* mRNAs (Treisman, 1985), and the transferrin

receptor mRNA (Müllner and Kühn, 1988). Sheiness *et al.* (1975) proposed that one of the most important roles of the poly(A) tail in mRNA is to protect mRNA from degradation, and there is some circumstantial evidence to support this. For example, c-*myc* and c-*fos* mRNA degradation is preceded by shortening of the poly(A)-tail (Treisman, 1985), whereas stable mRNAs, like globin, do not undergo poly(A) tail shortening. It has been suggested that histone mRNAs, which do not contain a poly(A) tail, form secondary structures at their 3'-end, which provide an alternative form of partial protection (Marzluff and Pandey, 1988). Also, there is some evidence that de-adenylation reduces the affinity of mRNA for ribosomes (Palatnik *et al.*, 1984); this might affect the translational activity (see section 1.9.6.).

Once sequences which accelerate RNA degradation become better characterised, care could be taken to ensure that they are not inadvertently incorporated into transcripts generated from transfected heterologous-genes. This might be achieved through use of synthetic genes with use different codons to those of the native gene (see section 1.9.6.).

1.9.6. Translation of mRNA.

Initiation of most eucaryotic-mRNA translation preferentially occurs at the AUG codon closest to the 5'-terminus (Kozak, 1983), especially when it is contained within the consensus sequence, GXXXAUGG (Kozak, 1986). These findings can be explained by a 'scanning ribosome' model, in which it is proposed that ribosomes can access initiation-sites only after first binding to the 5'-end of the mRNA. It is suggested that the ribosome then scans linearly along the RNA until an initiation site is recognised. However, recently discovered exceptions to this rule have been identified for Picornavirus RNAs, in which specific sequences allow initiation to occur at internal-initiation codons (Pelletier and Sonenberg, 1988; Jang *et al.*, 1988). For the majority of genes, therefore, if a promoter fragment is being used, which also contains a small piece of coding sequence from its native gene, it is important to remove the native translation-initiation codon to prevent synthesis of a fusion protein.

There are many examples of differential translation-rates of mRNAs. It has been recognised as a means of controlling expression from numerous cellular genes, particularly in embryonic development and

following heat-shock (see Darnell (1982) for a review). Also, translation control plays an important role in regulating expression of proteins produced in 'late' Adenovirus infection. In this situation, two elements are necessary to cause an increase in the translation rate; a tripartite-leader sequence which is present in the 5'-noncoding leader of late transcripts, and which is formed by splicing together of three small exons; and secondly, small, RNA-polymerase-III transcripts from the Adenovirus 'virus-associated' (VA) genes, VAI and VAII (Thummell *et al.*, 1983; Logan and Shenk, 1984). The VA RNAs are thought to increase translation by facilitating interaction of a 43S ribosomal-preinitiation-complex with tripartite-leader-containing mRNA, to form a 48S species (Schneider *et al.*, 1984). Kaufman (1985b) demonstrated that translational stimulation can also be conferred on hybrid transcripts, which contain sequences from exons 2 and 3 of the tripartite-leader, and a heterologous cDNA. It was suggested that such translation-control sequences might provide a general means of increasing the translation rate of any heterologous transcript. Hence, in this study, I attempted to reproduce such an effect in a myeloma cell (see section 3.8).

A mRNA may be poorly translated if the 5'-noncoding region contains excessive secondary-structure (Pelletier and Sonenberg, 1985). Hence, as a general rule, it is best to keep the 5-noncoding region of heterologous transcripts as small as possible (Cullen, 1986).

The genomes of different organisms have individual coding strategies, such that different codons may preferentially be used (Grantham, 1980). Different codons may also be preferentially used amongst genes of the same organism. This does not necessarily impose constraints on a protein sequence, because the degeneracy of the genetic code means that, for some amino-acids, more than one codon can code for the same amino-acid. This phenomenon has been best studied in bacteria. Here, it has been proposed that codon preference reflects a strategy to optimise translational kinetics under different growth conditions (Kurland, 1987). Codon preferences may affect the translation rate through differential stability of particular codon-anticodon interactions (Grosjean and Fiers, 1982). However, the concentration of tRNA isoacceptors has been shown to systematically correlate with codon frequencies in mRNA (Ikemura, 1981). Therefore, translation rates may also be affected by the availability of

specific tRNAs; if a mRNA contains codons for which the isoaccepting tRNAs are abundant and/or have high affinities for the codons, it may be translated more efficiently than a mRNA containing alternative synonymous-codons. This observation has been exploited in the expression of eucaryotic proteins in bacterial cells. Using synthetic genes, rather than the native gene, the codon composition can be optimised for translation in bacterial cells, resulting in increased levels of protein production. Codon usage is unlikely to be a major problem when expressing a heterologous mammaliam-protein in a mammalian host. Even so, synthetic genes might be used to contribute to high-level expression for other reasons, such as RNA stability (see section 1.9.5.). However, changing the codon usage of a gene might, in some cases, be detrimental to the production of a functional protein. For example, Purvis *et al.* (1987) hypothesised that sequences of rare codons may cause a pause in translation, which allows discrete protein domains to fold more efficiently. Hence, removal of such codons may prevent proper folding of a protein.

1.9.7. The secretory pathway.

The majority of proteins with potential therapeutic-use are, as discussed in section 1.1., blood products, which are, therefore, secreted from the producer cell. When producing such a protein in a heterologous host, it is advantageous to secrete it from the cell whenever possible. Indeed, a mammalian expression-host which could not secrete such proteins would usually not be considered suitable, because of three major reasons. Firstly, the protein is easier to purify from the medium than from cell lysates and, secondly, accumulation of a recombinant protein at high concentration inside the cell may have toxic effects. The third and most important reason, is that passage through the secretory pathway may be necessary to allow correct folding of the protein, and for post-translational modifications on the protein to take place (see section 1.10.), without which the protein may be non-functional.

For a protein to be secreted, it must enter into the endoplasmic reticulum (ER) lumen, then be transported to the golgi, and then be transported from the golgi to the cell surface. The protein is transferred between the cell compartments within membrane-bound vesicles. The entire

network is known as the secretory pathway (Palade, 1975). The only rational approach presently available for targetting a heterologous protein towards secretion, is to include a signal-sequence in the nascent protein, which should ensure delivery into the ER lumen (see section 1.10.1). Proteins may then be secreted through the secretory pathway, but the rules governing how a protein becomes secreted as opposed to being targetted to an intracellular destination, from within the secretory apparatus, are not well understood. The fact that the majority of therapeutic proteins are normally secreted, yet may not be secreted or are secreted inefficiently when produced in a heterologous host, implies there is cell-specificity in the factor(s) which affect secretion. Wieland *et al.* (1987) have shown that a tripeptide, presumably containing no destination signals, is efficiently secreted from HepG2 cells, following introduction into the ER. This suggests that no special features or signals are required for rapid and efficient transport of a protein from the ER lumen to the cell surface. Hence, proteins retained in the ER and golgi compartments are thought to be marked by specific retention-signals, for which there must be some receptor-based machinery to retain them in place against the bulk-flow (reviewed by Pfeffer and Rothman, 1987). One possible receptor is the BiP protein, which is found in the ER (see section 4.5.). Retention signals in proteins have, however, remained elusive. This may be because the signals are composed of regions of a protein called 'signal patches' (Pfeffer and Rothman, 1987), which are formed from noncontiguous regions of the polypeptide brought together during folding and are, therefore, conformation dependent. If proteins are retained in the secretory pathway for a considerable period of time, relative to their half-life, the levels of secretion will be severely affected. Therefore, it is desirable that the secretion of a heterologous protein be as rapid as possible. This requires the choice of, or the screening for, a host-cell which secretes the protein efficiently.

1.10. Post-translational modifications in eucaryotic cells.

As discussed in several preceding sections, most of the proteins with potential therapeutic-use are secretory proteins. When such proteins pass through the secretory pathway, they usually undergo post-translational

modifications. In order to ensure a protein produced in a heterologous host has the same properties as the native protein, it may be important that the host cell can reproduce these modifications. The various types of post-translational modifications, and the effects they may have on proteins, are described in the following sections.

1.10.1. Signal-sequence cleavage.

In order for secretion to be possible, proteins must be directed to the ER lumen. The ER is the first organelle of the secretory pathway (see section 1.9.7.). Targetting of a protein into the ER lumen occurs during protein synthesis (co-translationally). It is dependent on the presence of a signal-sequence, which is usually located at the NH_2-terminus of the nascent protein. The sequence consists of 16-26 amino-acid residues and has a polar, basic NH_2-terminus. and a central, apolar domain (von Heijne, 1983; Perlman and Halvorson, 1983). Usually, the signal-sequence is first recognised and bound by a signal-recognition particle (SRP) as it emerges from the ribosome, and this causes an arrest of translation (Walter and Blobel, 1981). The SRP then interacts with a SRP receptor ('docking protein'), which is situated in the ER membrane (Meyer *et al.*, 1982; Gilmore *et al.*, 1982). The signal-sequence is then transferred to a signal-sequence receptor (SSR), which is also situated in the ER membrane (Wiedmann *et al.*, 1987; Walter, 1987). It is believed this allows resumption of translation, and as the nascent chain elongates, the protein is translocated across the ER membrane. During this process the signal-sequence is usually proteolytically removed by a signal peptidase.

1.10.2. Other proteolytic cleavages.

In addition to the proteolytic cleavage of the signal-sequence, the highly specific removal of additional sequences from some proteins is necessary to allow biological function. For example, many proteolytic enzymes are produced as precursors known as zymogens, which are converted to the active protein by proteolytic removal of a 'pro' sequence. In some cases, such as in the production of the proteolytic enzyme, renin (Catanzaro *et al.*, 1983), this cleavage occurs in secretory vesicles prior

to secretion from the cell (see section 1.11.). In contrast, the processing and activation of some secretory proteins, such as blood-clotting factors and digestive enzymes, occurs extracellularly (reviewed by Neurath and Walsh, 1976). Other examples of intracellular proteolytic-cleavage include a number of polypeptide hormones such as adrenocorticotropic hormone, oxytocin and vasopressin, which arise from the proteolytic cleavage of large precursor-proteins (Richter, 1983). Also, inactive pro-insulin is converted to the active two-chain (A and B chains) form by proteolytic cleavage (Chan *et al.*, 1979).

1.10.3. Post-translational additions.

One of the most characteristic features of eucaryotic proteins is the way in which they become chemically modified following their synthesis. For secretory proteins, these modifications usually occur within the organelles of the secretory pathway, particularly the ER and the golgi. All modifications are likely to alter the characteristics of the protein in some way.

1.10.3.1. Glycosylation.

Glycosylation is the most common and best characterised post-translational chemical modification of proteins. It involves the addition of a polysaccharide side-chain to selected asparagine (N-linked), or serine and threonine (O-linked) residues. O-linked glycosylation appears to occur primarily in the golgi, where glycosyl transferases add sugar residues one by one. In contrast, N-linked oligosaccharides are synthesised by addition of a common mannose-rich precursor oligosaccharide, from a dolichol-pyrophosphoryl oligosaccharide molecule, in the ER. This is followed by processing and modification of the carbohydrate chain to a more complex structure, in both the ER and golgi.

The fine structure of oligosaccharides on the same polypeptide have been shown to vary with host cell, transformation state and establishment in culture (Kaetzel *et al.*, 1985). Also, Sheares and Robbins (1986) showed expression from the chicken-ovalbumin gene in mouse L-cells results in production of recombinant ovalbumin, in which the same

asparagine residue is glycosylated as in native chicken-ovalbumin, but in which the pattern of individual sugar residues in the oligosaccharide chain is novel. Native ovalbumin, however, also displays heterogeneity with respect to the oligosaccharide chain; at least nine basic-structures have been characterised. Therefore, this may suggest a precise glycosylation pattern is not essential for the protein to be functional. In fact, the function of the carbohydrate moieties of most proteins has not been elucidated, and in some cases the absence of glycosylation has no discernable effect (Edge and Camble, 1984; Liang *et al.*, 1985). There are few reports of glycosylation being directly necessary for functional activity. A study by Sairam and Bhargavic (1985) suggests that carbohydrate is necessary for transduction of a biological signal in a gonadotropic-hormone receptor (Sairam and Bhargavic, 1985). Also, a recent study suggested that when the heavily-glycosylated protein, erythropoietin, is expressed in different mammaliam expression-hosts, the biological activities of the proteins differ due to different carbohydrate attachments (Goto *et al.*, 1988). However, differences between glycosylated and non-glycosylated forms of a protein are more likely to be reflected in the general characteristics described below.

A carbohydrate chain can play an important role in the antigenicity of the protein. It may contribute the most important epitope of the protein. Therefore, for example, if a protein vaccine is produced in a non-glycosylating host cell, antibodies will not be raised, in the vaccine recipient, against the most appropriate feature of the protein. Consequently, the antiserum of the recipient may have poor affinity for the antigen. Equally, the absence of the carbohydrate chain may unmask additional antigenic-determinants, which could be recognised as foreign, and prove immunogenic when a protein is administered to a recipient, even though the protein is normally present in the recipient, albeit in a different form.

The carbohydrate sometimes provides protection against proteolytic enzymes. This may be important during purification of the protein, and may also increase the plasma half-life of the protein. Also, the carbohydrate may be recognised by receptors on cell surfaces, which play an important role in determining the clearance rate of proteins from the plasma. For example, receptors on the surface of liver cells have an

affinity for mannose residues of glycoproteins. Gross *et al.* (1988) demonstrated that N-linked oligosaccharides are important for plasma survival of liver secretory-glycoproteins, and that unphysiologically-glycosylated forms are cleared by different mechanisms.

The N-linked carbohydrate chains often contain highly charged residues such as sialic acid, which may increase the solubility of a protein. This may be an important consideration when formulating therapeutic agents at clinically relevant concentrations.

The carbohydrate residues added will inevitably affect the conformation of the protein, which may also influence the above discussed properties of antigenicity, stability and solubility. In addition, a specific conformation may be necessary for biological activity of the protein. There is evidence that early glycosylation events on the nascent polypeptide, during translation, actually influence the folding of the chain. The best studied example of this phenomenon is the vesicular-stomatitis virus G-protein, which is found on the surface of both the virus and cells infected with the virus. Studies suggest that the cotranslational attachment of the precursor oligosaccharide to the G-protein can serve as a nucleation centre which facilitates protein folding (Leavitt *et al.*, 1977).

1.10.3.2. Disulphide bonds.

Disulphide bonding between two cysteine residues is believed to be one of the most important stabilising-forces in the secondary structure of a protein. Disulphide bonds are generally confined to secretory proteins and membrane proteins. Intracellular proteins are unlikely to contain disulphide bonds, because of the reducing nature of the intracellular environment. The formation of the bonds often occurs immediately following entry of the polypeptide to the lumenal side of the ER. It may occur sequentially as the protein is growing and, therefore, play an important role in the native folding of the proteins (Bergman and Kuehl, 1974). The reaction is catalysed by the enzyme, protein disulphide-isomerase (Freedman *et al.*, 1988)

1.10.3.3. Addition of lipid.

A surprisingly large number of eucaryotic proteins contain covalently-bound lipid (for a review see Sefton and Buss, 1987). A diverse class of proteins are found to contain the 16-carbon saturated fatty-acid, palmitic acid, linked to cysteine. Some soluble proteins inside the cell contain the 14-carbon saturated fatty-acid, myristic acid, at their amino-termini, but its function is not well understood. Cell-surface proteins may be anchored to the outer cell-surface by a complex, glycosylated phospholipid. It is believed that the binding of membranes is the most important function of lipid added to proteins. The existence of and importance of lipid in secretory proteins has not been extensively studied.

1.10.3.4. Other post-translational additions.

In a recent count, more than 150 different post-translational modifications to proteins had been identified (Freedman, 1988). It is likely that as more proteins are characterised in detail, more novel modifications will become apparent.

The amino-termini of many proteins are acylated, commonly with an acetyl group (Narita, 1975), which seems to make them more resistant to degradation. Amidation has been shown to occur on the carboxy-terminal amino-acid of some proteins, especially pituitary hormones (Lowry and Chadwick, 1970), such as calcitonin. Selenization is required to generate the active form of glutathione peroxidase (see Spier, 1986). Also, γ-carboxylation (Senflo *et al.*, 1974; Nelsestuen *et al.*, 1974) and β-hydroxylation (see Furie and Furie, 1988) modifications occur on several blood-clotting factors, and are involved in generating the active state. For example, a sequence of glutamic-acid residues, at the amino-terminal ends of factors VII, IX and X, becomes γ-carboxylated. It has been shown that Ca^{2+} binding to this region of the proteins allows the factors to assemble on membrane surfaces, and take part in reaction complexes (Fujikawa *et al.*, 1974). If factor-IX is expressed in cells that do not reproduce this modification, the protein shows very poor activity (M.Courtney and J.-P. Lecoq, personal communication). In addition, many recently-discovered enzymes phosphorylate proteins on either tyrosine, threonine or serine residues. However, the majority of proteins on which phosphorylation has been identified are intracellular, such as metabolic

enzymes, and proteins implicated in the growth-control mechanisms of the cell. Hence, it is probable that phosphorylation is not relevant to the majority of therapeutic proteins, which are secreted.

1.10.4. Oligomerisation.

Oligomeric proteins are proteins which consist of more than one polypeptide chain. Such proteins can be produced by cleavage of a single precursor (e.g. insulin; see section 1.10.2.), but in the majority of cases, they are produced by association of separate polypeptide chains (oligomerisation). For secretory proteins, oligomerisation usually occurs within the secretory pathway. As will be discussed in section 4.5., the immunoglobulin heavy and light chains associate within the secretory pathway, and in some cases the chains are unable to pass through the secretory pathway unless they do associate.

1.11. Myeloma cells: a potential mammalian cell expression-host.

Myeloma cells are the neoplastic counterparts of plasma cells. Plasma cells are terminally-differentiated B-lymphoid cells, which normally express and secrete large amounts of immunoglobulin (Ig). As discussed in section 1.3.5., there is a need for a larger number of mammalian cell expression-systems than is presently available. The aim of this investigation was to develop and assess the use of a mouse, myeloma cell-line as an expression host for heterologous eucaryotic-proteins. There are three major reasons why myeloma cells are considered an attractive potential expression-host for heterologous proteins, as discussed below.

Firstly, myeloma cells are likely to secrete proteins efficiently. Secretory cells fall into 2 major classes, namely constitutive or regulated secretory-cells, as emphasised by Tartakoff and Vassali (1978). Constitutive cells lack short-term control of the secretion rate; small golgi-derived non-clathrin-coated vesicles, carrying secretory proteins, pass directly to the cell surface in a stimulus-independent fashion, where they fuse with the plasma membrane to release their contents by exocytosis (Orci *et al*, 1986). In contrast, in regulated secretory cells, the golgi elements give rise to immature condensing vesicles, that

convert to storage vesicles containing a partial clathrin coat (Orci *et al.*, 1984). These vesicles fuse with the cell surface only upon receipt of an extracellular signal (hormone or neurotransmitter). Myeloma cells are a prime example of a constitutive secretory-cell. Therefore, they do not contain a large intracellular pool of Ig, even though they synthesise large amounts of Ig; all the Ig synthesised is rapidly secreted and Ig accumulates at approximately 10μg/ml in the medium of a normal myeloma-cell culture. In the myeloma cell, MPC-11, as much as 20% of total protein-synthesis may be committed to Ig chains, and the secreted antibody (IgG_{2b},κ) can represent almost 100% of all secreted protein (Laskov and Scharff, 1970; calculations are based on the assumption that the valine, threonine and leucine content of Ig chains is typical of total cell-protein). In order to accommodate the large amounts of secreted Ig, the secretory network of a myeloma cell is extensive. This makes it less likely that the amount of secretory apparatus will limit the secretion levels of heterologous proteins. In addition, if heterologous proteins can be secreted at levels similar to Ig, the fraction of contaminating proteins in the medium will be small. This should make purification of the recombinant product a simple process. As yet, it is unclear whether constitutive secretory-cells, such as myelomas, are able to carry out the proteolytic-processing steps that may occur, within secretory vesicles (see section 1.10.2.), on proteins which are normally secreted by a regulated pathway (Kelly, 1985).

A second advantage of myeloma cells is that they are non-adherent, and grow well in suspension. The production of monoclonal antibodies from hybridomas (products of fused myelomas and spleen lymphocytes) in large-scale suspension culture is a well-developed process, and much of this expertise is likely to be applicable to the growth of myelomas. For example, the homogeneous agitated-bioreactor system, described by Reuveny *et al.* (1986), might be suitable. In addition, serum-free media, developed for the growth of hybridomas, might also be used for growth of myelomas, which could reduce costs.

Finally, the molecular biology of Ig expression has been well studied, and some of the expression elements responsible for high-level Ig expression in myeloma cells have been characterised (as is discussed in section 1.12.). Such information, together with the availability of the

expression elements, is likely to be useful for obtaining high-level expression of the products of heterologous genes in myeloma cells.

1.12. Immunoglobulin heavy-chain (IgH) expression.

The Ig genes are some of the most well-studied genes with respect to control of gene expression. The following sections summarise the current understanding of the control of IgH expression. They illustrate the complexity of expression control, and also how expression can be controlled at both transcriptional and post-transcriptional levels. It therefore serves as a specific example of some of the principles discussed in section 1.9. Also, as alluded to above, the knowledge on, and availability of Ig expression-elements was invaluable, in this study, towards the development of an expression system for a myeloma cell. Since myeloma cells express Ig at high levels, it was rationalised that the elements contributing to high-level Ig expression might also allow high-level expression from heterologous genes.

1.12.1. Quantitative changes in IgH expression.

In the course of B-lymphoid-cell development, an IgH(μ) gene is assembled by fusion of germ-line DNA segments during a series of recombination events. One, of more than 100, variable(V_H)-region gene segments becomes fused, at its 3'-end, with a diversity(D)-region segment, which in turn is fused, at its 3'-end, with a joining(J)-region segment (Early *et al.*, 1980; Tonegawa, 1983). This string of segments (which is, together with the leader (L) sequence, referred to as the V_H-region in this study) then becomes fused, at its 3'-end, with the μ-constant-region (Cμ) sequence, which consists of a large (major) intron, Cμ-region exons and associated introns, and downstream sequences.

B-cells, containing the rearranged μ-gene, synthesise functional μ-transcripts from a promoter 5' to the V_H-region (Clarke *et al.*, 1982). In contrast, prior to recombination, the only transcripts generated from the Cμ region are sterile, and are initiated from secondary or cryptic promoters in the 5'-flanking region (Nelson *et al.*, 1983). The B-cells mature into virgin B-lymphocytes. On encountering antigen, these are

activated (Melchers *et al.*, 1984) and may give rise to the terminally-differentiated antibody-secreting plasma cells. The antibody secreted may contain the original class of heavy chain (μ chain), or another class of heavy chain, which results from a further recombination event called 'switching', in which the Cμ region of the gene is substituted by another class of Cμ region. The plasma cells accumulate Ig mRNA at a 30-100 - fold higher concentration than in B-cells (Perry and Kelley, 1979). Most of this dramatic increase is due to post-transcriptional regulation, because the rate of IgH-gene transcription in plasma cells is only accelerated 2-5 - fold over the rate in B-cells (Mather *et al.*, 1984; Gerster *et al.*, 1986).

1.12.2. Elements controlling IgH expression.

Many of the elements important for expression from IgH genes were identified using gene-transfer experiments. In these experiments, cloned, rearranged wild-type or *in-vitro*-mutated IgH genes were introduced into terminally differentiated lymphoid, and non-lymphoid cells. The levels of Ig expression obtained in these experiments allowed identification of the first cellular-enhancer (Banerji *et al.*; Gillies *et al.*; Neuberger, all 1983). The IgH enhancer is located in the major intron, between the J- and C-region exons, of IgH genes. The enhancer displays cell-specificity in that it only augments transcription in cells of lymphoid lineage. The IgH promoter is the second essential element for high level-transcription from IgH genes, and also shows some cell-type specificity (Grosschedl *et al.*; Foster *et al.*; Mason *et al.*, both 1985). The IgH enhancer and promoter are reported to show preferential synergy in myeloma cells, because the IgH enhancer potentiates transcription from the IgH promoter more effectively than it does on heterologous promoters (Garcia *et al.*, 1986). It has been proposed that the IgH(μ) gene contains at least a third regulatory element because, when isolated from the enhancer and promoter, the remaining intragenic sequences still confer lymphoid-cell specificity to heterologous transcription-units (Grosschedl and Baltimore, 1985). The location of this regulatory element has not yet been identified. However, since most of the cell-stage-specific regulation of the μ-mRNA level is post-transcriptional, the intragenic sequences are likely to regulate pre-mRNA splicing, RNA transport or mRNA stability. Also, Gregor *et al.* (1986) have shown that

deletion of a 4Kb DNA-fragment immediately downstream of the α gene (IgA heavy-chain gene) decreases the transcription rate of the gene. This suggests the presence of another regulatory element associated with IgH genes.

In lymphoid cells, the enhancer is probably under positive control, such that binding of *trans*-acting factors causes an increase in transcription. *In vivo* and *in vitro* binding-studies, using footprinting and gel-retardation (mobility-shift) assays, have produced direct physical evidence for binding of cellular factors to the IgH enhancer and promoter. For example, Ephrussi *et al.* and Church *et al.* (both 1985) demonstrated binding to three enhancer-motifs *in vitro*, and an additional two sites *in vivo*. Different factors have been shown to recognise the different *cis*-acting elements (Weinberger *et al.*, 1986). The most interesting *cis*-acting sequence is a highly conserved octamer, which is common to all Ig genes, and is found in both the promoter and the enhancer of the IgH gene (reviewed by Dynan, 1987). Some of the *trans*-acting factors also interact with other enhancer elements which do not show lymphoid-cell specificity (Scholkat *et al.*, 1986), supporting the idea that a combination of cell-specific and general factors are necessary for activation of enhancers (Voss *et al.*, 1986). It also appears that some of the cell-type specificity of the IgH enhancer results from the binding of negative regulatory-factors in non-lymphoid cells (Kadesch *et al.*, 1986; Imler *et al.*, 1987). The removal of this *cis*-acting 'repressor' sequence, which normally binds to the negative *trans*-acting factor(s), from the IgH enhancer, allows some activity of the enhancer in fibroblasts, a cell type in which it is normally inactive (Weinberger *et al.*, 1988). In common with many other enhancers, the IgH enhancer shows considerable redundancy, because any one of the *trans*-acting factor binding-sites can be deleted without a large effect on enhancer activity (Gerster *et al.*, 1987).

Interestingly, there are spontaneous myeloma-cell mutants, in which the supposedly important enhancer is deleted from the IgH gene, yet the encoded IgH-chain is expressed at normal levels (Klein *et al.*, 1984; Eckhardt and Birstein, 1985; Wabl and Burrows, 1985). This apparent contradiction was resolved when the mutant genes were cloned, and then reintroduced into myeloma cells. IgH was not expressed from these transfected mutant-genes (Klein *et al.*, 1985; Zaller and Eckhardt, 1985).

This suggests that the effect of the enhancer is temporal, possibly being responsible for the initial assembly of a stable transcription-enabling complex, that can be maintained in the absence of the enhancer.

As discussed in section 1.12.1., a recombination event during early lymphoid-cell development results in juxtapositioning of the enhancer downstream of, and nearby, a previously silent V_H-region promoter. Considering the importance of the enhancer in ensuring high-level expression from transfected IgH-genes, it was long regarded that the juxtapositioning of the enhancer nearby the V_H-region promoter during the recombination event was responsible for activation of the V_H-region promoter, to allow high-level production of complete μ-transcripts. This was supported by the fact that the germ-line V_H-regions are not transcribed in myeloma cells. However, a demonstration that these unrearranged V_H-regions are transcriptionally active in the early stages of B-lymphocyte development (Yancopoulus and Alt, 1985) suggests that the control of Ig expression is in fact far more intricate.

1.12.3. Qualitative changes in IgH expression.

In addition to the quantitative changes in Ig expression during B-lymphocyte differentiation, there are also qualitative changes. I have already described 'switching', in which the constant region of the gene may be replaced by an alternative class of constant region. However, even if the class of IgH gene is not altered, qualitative changes in the IgH chain produced still occur during B-lymphocyte development, as discussed below.

Two forms of the IgH(μ) chain can be produced by a single lymphoid cell. Their protein sequences differ at the C-terminal; a membrane associated form (μ_m) contains a hydrophobic anchor-segment, whereas the secreted form (μ_s) contains a hydrophilic-segment instead. The two proteins are encoded by two different mRNA species that also differ at their 3'-ends, but are encoded by a single rearranged-IgH-gene (Alt *et al.*, 1980; Early *et al.*, 1980). μ_m mRNA is predominant during early stages of development, but in the plasma cell, the ratio is altered in favour of μ_s mRNA, which leads to a 10-100 - fold excess of the secreted chain (Tyler *et al.*, 1982). The production of μ_m mRNA entails the use of a distal poly(A)-site, and a splice that eliminates the μ_s terminus. The production of μ_s

entails use of a proximal poly(A) site that eliminates the 3'-splice junction. The positions of the poly(A) sites and the membrane region exons in an IgH(μ) gene are illustrated in the diagram of the plasmid, pSV-Vμ1 (see figure 4.4.).

Three models have been proposed to account for the regulation of the ratio between the two forms of IgH mRNA. One model suggests that it is regulated through use of alternative transcription-termination sites. Hence, if termination preferentially occurs after the sequences coding for μ_s but before the sequences coding for μ_m, only μ_s mRNA can be produced. However, this mechanism seems to operate in only a small number of mature B-cells (Kelley and Perry, 1986; Galli *et al.*, 1987). In the majority of mature B-cells (e.g. plasmacytomas PC 3741 and 104.76), the transcription rate across the μ_s and μ_m encoding sequences is equimolar, regardless of the μ_s:μ_m mRNA ratio (Yuan and Tucker, 1984; Milcareck and Hall, 1985; Kelley and Perry, 1986; Reuther *et al.*, 1986). In these cells, therefore, one of two further models is likely to be more accurate. The regulation could work at either the level of differential cleavage, so as to affect the rate of μ_s-mRNA production, or by differential splicing, so as to regulate the incorporation of the membrane exons into the transcript. It is not possible to distinguish unequivocally between the two models; if cleavage occurs at the μ_s poly(A)-site, splicing is precluded, and vice versa. The μ_m:μ_s ratio depends on the distance between the two poly(A) sites (Peterson and Perry, 1986). This suggests that differential processing is a kinetically-controlled event, and may depend on the availability of *trans*-acting factors which changes during development (Gregor *et al.*, 1986).

In some cases, on differentiation into a plasma cell, the membrane-associated Ig disappears completely from the cell surface. Although progressive disappearance of μ_m mRNA seems to be the predominant explanation, it has been reported that μ_m mRNA shows reduced translational efficiency in secretory cells, and also the membrane-associated Ig appears to be retained in the ER and becomes degraded before it reaches the cell surface (Sitia *et al.*, 1987).

1.13. Aim of the investigation.

The aim of this study was to develop and assess the use of myeloma cells as an expression host for the secretion of large amounts of cloned animal-gene products. The mouse, myeloma, cell-line, J558L (Oi *et al.*, 1983) was used as a model expression-host. This is because at the outset of the project, it was the most-easily-transfected myeloma cell-line available (see section 1.4.). J558L is a spontaneous IgH-chain-loss variant, derived from the cell-line, J558 [α,λ; anti-α 1-3 dextran] (Lumblad *et al.*, 1972). The laboratory as a whole wishes to assess the versatility of myeloma cells for the expression of heterologous proteins. Also, a mammalian expression-system is required as as a tool for basic-research purposes (see chapter 6). However, before such work could be attempted, it was necessary to develop a highly-efficient expression system. The thesis describes the various approaches I pursued in order to optimise expression from the chicken-lysozyme gene which acted as a model system. The study demonstrated that expression and secretion of a heterologous protein can be obtained, and at levels similar to that of normal Ig-expression from myeloma cells.

CHAPTER 2: MATERIALS AND METHODS.

2.1 Materials and plasmids.

2.1.1. Source of materials.

The suppliers of enzymes, antibiotics, chemicals and media are listed below. All chemicals were of analytical grade unless stated otherwise.

Aldrich Chemical Company, Gillingham, Dorset, England.
Rubidium chloride.

Amersham International plc, Little Chalfont, England.
Radionucleotides α-^{35}S-dATP, α-^{32}P-dATP and ^{35}S-methionine.

BDH Biochemicals, Poole, England.
Amberlite-mixed bed resin (MB1), Brij 35, bromophenol blue, glycine, 8-hydroxyquinoline, LiCl, N,N'-dimethylformamide, NP40, PEG 1500, xylene cyanole FF.

Bio-Rad Laboratories, Watford, England.
Ammonium persulphate.

Boehringer Corporation Ltd., Lewes, England.
CIP, oligo(dT)-cellulose.

Bethesda Research Laboratories, Bethesda, Maryland, USA.
CsCl (optical grade), low melting point agarose (ultra-pure).

Corning Glass Works, Corning, N.Y., USA.
75cm^2 tissue culture flasks, 50 ml conical tubes with screw cap.

Difco Corporation, Detroit, USA.
Bacto-agar, bacto-tryptone, bacto-yeast extract.

Flow Laboratories, Irvine, Scotland.

Eagle's modified minimum essential medium (with Earle's salts and 2g/l sodium bicarbonate, and without methionine and glutamine), Iscove's medium (supplemented with transferrin, soybean lipid, glutamine, 25mM Hepes buffer and 3.024g/l sodium bicarbonate).

Fisons plc., Loughborough, England.

Boric acid, butan-1-ol, $CaCl_2$, chloroform, D-glucose, dimethyldichlorosilane solution, EDTA, ethanol (IMS and absolute), formaldehyde, glacial acetic acid, glycerol, glycine, HCl, iso-amyl alcohol, isopropanol, KCl, KH_2PO_4, methanol, $MgCl_2$, $MgSO_4$, Na acetate, Na azide, Na citrate, NaCl, Na_2HPO_4, NaOH, Na salicylate, NH_4 acetate, NH_4Cl, $(NH_4)_2SO_4$, PEG 6000 (SLR quality), phenol, scintillation cocktails, SDS, sucrose.

Gibco-BRL plc., Paisley, Scotland.

AMV reverse transcriptase, Bal 31, DMEM, L-glutamine solution (for cell culture), Klenow, penicillin / streptomycin solution (for cell culture), all restriction enzymes, T4 DNA ligase, T4 DNA polymerase.

Lilly Research Centre, Windlesham, Surrey, England.

Mycophenolic acid (lot P73686).

Miles Laboratories Ltd., Slough, England.

Agarose, Freund's complete adjuvant, rabbit anti-mouse IgG antibody, rabbit anti-chicken ovalbumin antibody.

Nunc (A/S), Roskilde, Denmark.

96 well microtitre plates, 24 well limbro plates, $25cm^2$ tissue culture flasks, $150cm^2$ tissue culture flasks.

Scientific Industries International inc. (UK) Ltd., Loughborough, England.

Dialysis membrane.

Sera Lab., Crawley Down, Sussex, England.

Foetal calf serum, donor horse serum.

Serva, Heidelberg, W.Germany.

Acrylamide, N,N'-methylene-bisacrylamide, urea.

Sigma Chemical Company Ltd., Poole, England.

Agarose (type II-medium EEO), ampicillin (for bacterial work), ATP, β-mercaptoethanol, BSA, $CdSO_4$, chloramphenicol, Coomassie brilliant blue G, deoxycholic acid (Na salt), diethylpyrocarbonate solution, DNase I, dNTPs (dATP, dCTP, dGTP, dTTP), ddNTPs (ddATP, ddCTP, ddGTP, ddTTP), DTT, EGTA, EtBr, ficoll type 400, formamide, heparin (Na salt, low mol.wt.), Hepes buffer, herring testes DNA, hypoxanthine, IPTG, kanamycin, chicken egg lysozyme, $MnCl_2$, Mops buffer, pancreatic RNase (type X-A), Pipes buffer, polyvinylpyrrolidone-40, spermidine, standard protein mol.wt. markers, TEMED, tetracycline, thiamine, Tris buffer, trypan blue dye, xanthine (cell culture reagent), X-gal.

Whatman Biochemicals Ltd., Maidstone, Kent.

3MM chromatography paper, No. 1 filter paper.

2.1.2. Source of plasmids.

Name of plasmid.	Sit of referal	Source of plasmid.
M13mp9	section 2.10.1.1.	Vieira and Messing (1982)
M13mp18	section 2.10.1.1.	Yanisch-Perron *et al.* (1985)
M13mp9-gpt	section 2.10.1.1.	this work
M13mp9-5'gpt	section 2.10.1.1.	this work
M13mp18-ova	section 2.10.1.1.	this work
M13mp18-lys	section 2.10.1.1.	this work
pβ5'SVBgII	section 4.7.5.	Grosveld *et al.* (1982)
pBR322	section 4.8.1.	Bolivar *et al.* (1982)
pBR322-ΔH3	section 4.8.1.	this work
pCMV-gpt	section 3.4.	this work
pCMV-Egpt	section 3.3.1.	this work
pD20	section 3.8.	Kaufman (1985)
pΔTK-Egpt	section 3.3.1.	this work
pΔTK-Egpt-ΔH3.,ΔB.	section 4.7.1.	this work

pΔTK-Egpt-CMV	section 4.7.1.	this work
pΔTK-E'gpt	section 3.3.1..	this work
pΔTK-gpt	section 3.4.	this work
p-Egpt	section 3.5.	this work
pElysE(t)	section 4.8.1.	this work
pElys(t)	section 4.8.1.	this work
pEovaE(t)	section 4.8.1.	this work
pEova(t)	section 4.8.1.	this work
p-gpt	section 3.5.	this work
pIg-Egpt	section 3.3.1.	this work
pIg-gpt	section 3.4.	this work
plys-βglo	section 4.7.5.	this work
plysE(t)	section 4.8.1.	this work
$plys^+$(d)	section 4.7.1.	this work
$plys^+$(d)-ΔH3	section 4.8.1.	this work
$plys^-$(d)	section 4.7.1.	this work
plys-ISV40	section 4.7.5.	this work
plys-SV40	section 4.7.5.	this work
plys-SV40-3'$H2A^+$	section 5.3.	this work
plys-SV40-3'$H2A^-$	section 5.3.	this work
plys-TK	section 4.7.5.	this work
plysμ	section 4.9.1.	this work
plysμ2	section 4.9.1.	this work
plysμA	section 4.9.1.	this work
plysμB	section 4.9.1.	this work
plysμC	section 4.9.1.	this work
plysμD	section 4.9.1.	this work
plysμA-Δ1	section 4.9.1.	this work
plysμA-Δ2	section 4.9.1.	this work
plysμA-Δ3	section 4.9.1.	this work
plysμA-Δ4	section 4.9.1.	this work
pMHVA	section 3.8.	Harrison, T.M. (unpublished)
plysμA-Δ5	section 4.9.1.	this work
pMLP-CAT	section 3.3.1.	Gift of M. Matthews
pML-Egpt	section 3.3.1.	this work
pML-gpt	section 3.4.	this work

pMLT-Egpt	section 3.3.1.	this work
pMLT-Egpt-VA	section 3.8.	this work
pMMTV-Egpt	section 3.3.1.	this work
pMMTV-gpt	section 3.4.	this work
pMTEB1A	section 3.3.1.	Zeller *et al.* (1986)
pMT-Egpt	section 3.3.1.	this work
pMT-gpt	section 3.4.	this work
pova(d)	section 4.7.1.	this work
povaE(t)	section 4.8.1.	this work
pRSV-CAT	section 3.4.	Gorman *et al.* (1982)
pSV-Egpt	section 3.3.1.	this work
pSV-gpt	section 3.4.	this work
pSV_2gpt	section 3.3.1.	this work but also constructed by Mulligan and Berg (1980)
pSV2gpt-ΔR1	section 3.3.1.	this work
pSV_2gptE	section 5.2.1.	this work
pSV_2gptE'	section 5.2.1.	this work
pSV_2gpt-3'H2A$^+$	section 5.3.	this work
pSV_2gpt-3'H2A$^-$	section 5.3.	this work
pSVMdhfr	section 3.3.1..	Lee *et al.* (1981)
pSV_1ov	section 4.2.3.	Krieg *et al.* (1984)
pSV_2ov	section 4.2.4.	Krieg *et al.* (1984)
pSV-Vμ1	section 4.3.	Neuberger (1983)
pTK	section 3.3.1.	this work
pTK1	section 4.7.5..	Colbere-Garapin *et al.* (1979)
pTK-E(a)	section 3.3.1.	this work
pTK-E(b')	section 3.3.1.	this work
pTK-E(b')gpt	section 3.3.1.	this work
pTK-E(b')ova	section 4.2.2.	this work
pTK_2ov	section 4.2.2.	Krieg *et al.* (1984)
pTK_2ov-ΔH3	section 4.2.2.	this work
pTK_2lys	section 4.7.1.	Krieg *et al.* (1984)
pUC7-CMV	section 3.3.1.	Akrigg *et al.* (1985)
pUC7-CMV-ΔSst	section 4.7.1.	this work
pUC8	section 3.3.1.	Vieira and Messing (1982)
pUC8-3'H2A$^+$	section 5.3.	this work

pUC8-3'H2A⁻	section 5.3.	this work
pUC8-Ig	section 3.3.1.	this work
pUC8-MMTV	section 3.3.1.	this work
pUC8-SV40	section 3.3.1	this work
pUC8-ULVDJ$_2$	section 3.3.1.	this work
pUC18	section 4.7.5.	Yanisch-Perron *et al.* (1985)
pμ2	section 4.9.1.	S.Munson (lab. colleague)
pVμ1	section 4.9.1.	J.Ridge (lab. colleague)
pXTKVδ1	section 3.3.1.	Neuberger (1983)
pXTKVδ6	section 3.3.1.	Neuberger (1983)

2.2. Myeloma-cell culture.

2.2.1. General handling and media.

The mouse, myeloma cell-lines used in this study were J558L (Oi *et al.* 1983), X63.Ag8 (Köhler and Milstein, 1975) and NSO (Galfrè and Milstein, 1981). All cell lines were routinely maintained in maintenance medium (Dulbecco's modified medium containing 4.5g/l glucose and without sodium pyruvate (DMEM), supplemented with 10% (v/v) horse serum (HS) and with antibiotics penicillin (50 I.U./ml) and streptomycin (50 μg/ml)). The concentrations of supplements used in maintenance medium were used in all other media, except where stated otherwise. Cultures were maintained at a cell density of between 3 x 10^5 and 2 x 10^6 cells/ml, in plastic 24 x 2ml well ('Linbro') plates or in 75cm^2 Sterilin tissue-culture flasks. The incubators were kept at 37°C, and were supplied with a humidified atmosphere of 10% CO_2 (v/v) in air.

Cells densities were determined using a Neubauer haemocytometer and a Vickers microscope. Accurate determinations of cell viability were performed using a trypan-dye exclusion test; 10μl of trypan-dye solution (0.2% (w/v) trypan blue in cPBS) was added to 90μl of a cell suspension, and cells not taking up the dye were regarded as viable.

In all experiments, myeloma cells were pelleted at 1,000 rpm in a MSE bench-centrifuge at room-temperature; if the cell suspension volume was greater than 5ml, Corning, 50ml orange-capped tubes were used, but for

smaller volumes, disposable Sterilin, 10ml, capped test-tubes were more convenient. Routine manipulations were carried out in a still air-hood, but potentially hazardous experiments, such as transfections, were performed in a Sterilgard laminar-flow air-hood. Aseptic technique was used at all times and surfaces were frequently swabbed with 70% (v/v) IMS.

2.2.2. Cloning in Agarose.

Agarose cloning was used to accurately determine the cloning efficiency. This is expressed as the percentage of cells able to proliferate and give rise to colonies. It was used to assess the toxic effects of compounds added to the medium, and also to screen various batches of sera for their ability to support cell-growth. Cells were diluted to a density of 20 cells/ml in a DMEM solution containing 2% (v/v) FCS and penicillin / streptomycin. 9ml of this solution was added to 1ml of a molten agarose-solution (1% (w/v)), which had been pre-incubated at 50°C. This agarose solution had been made by mixing equal volumes of 2% (w/v) agarose and DMEM / penicillin / streptomycin. The resultant suspension was poured into 10ml test-tubes, the tubes were capped, and then the agarose was allowed to solidify before putting the tubes in the incubator. The number of colonies produced was counted after 14 days.

2.2.3. Cloning / subcloning.

In order to ensure a cell population being analysed was monoclonal rather than heterogeneous, subcloning was performed. Cells were diluted to a cell density of 20 cells/ml, and 50µl aliquots of this suspension (average of one cell) were placed in wells of a 96-well microtitre plate. If the cells were fed with conditioned medium (partially spent medium from other cell-cultures), it was possible to produce colonies in more than 1/3 of the wells. Colonies were microscopically examined, to ensure they were monoclonal, before being picked.

2.2.4. Preservation of cell-lines.

Important cell-lines were preserved by freezing in liquid-nitrogen. Approximately 2 x 10^7 cells, which were undergoing log-phase growth, were pelleted. The cells were resuspended in 1ml of a DMEM solution containing 45% (v/v) FCS and 10% (v/v) DMSO, and then transferred to a 1.5ml Sarstedt freezing-vial. After standing on ice for 20 minutes, the vials were insulated by wrapping them in tissue-paper and then placing them in a polystyrene carton. The cells were allowed to freeze slowly by placing the carton in a -70°C freezer for 24 hours. The vials were then transferred to liquid-nitrogen for storage.

Cells were thawed out rapidly, by rotating the vial in a 37°C water-bath until all ice-crystals had melted. The cells were then diluted out of DMSO by adding 20 volumes of DMEM, and then pelleted. The cell pellet was resuspended in 'recovery' medium (medium used following transfections; see section 2.3.1.) at a range of cell densities (1 x 10^5 - 2 x 10^6 cells/ml).

2.3. Transfection of myeloma cells.

2.3.1. Spheroplast fusion.

Almost all transfections were carried out by spheroplast fusion using a method based on that described by Oi *et al.* (1983). *E. coli* cells, containing the plasmid to be introduced into the myeloma cell, were grown as an overnight culture in 10ml of LB-broth / ampicillin. This suspension was diluted 1:100 (v/v) in 50ml of LB-broth / ampicillin and grown for approximately 2½ hours at 37°C until the OD_{650} was 0.7 (accepted range: 0.6-0.8). A 100mg/ml stock-solution of chloramphenicol, dissolved in ethanol, was added, to give a final concentration of 0.2 mg/ml. Then the plasmid copy-number was allowed to amplify, by incubating for a further 16 hours. 25 ml of the resultant culture was centrifuged in a MSE-Chilspin centrifuge at 4,000 rpm, for 5 minutes at 4°C, and the supernatant was removed as completely as possible.

Spheroplasts were then prepared from the cell pellet. In order to produce stable spheroplasts, it was necessary to carry out all steps from this point onwards with ice-cold solutions. The cell pellet was resuspended in 0.25ml of a solution containing 50mM Tris.HCl (pH8.0) and 20% (w/v)

sucrose, and then 50µl of freshly prepared chicken-lysozyme solution (10 mg/ml) was added. After exactly 5 minutes, 125µl of 0.25M EDTA (pH8.0) was added. This was followed by careful addition, with stirring, of a solution containing 50mM Tris.HCl (pH8.0) and 9% (w/v) sucrose; 1ml was added over 1½ minutes, followed by 2.25 ml over the next 1½ minutes. This spheroplast suspension was further diluted with 3.75ml of a DMEM solution containing 7% (w/v) sucrose, 15mM $MgCl_2$, 15mM EGTA (pH7.5) and 2 x 10^{-4}% (w/v) DNAse-1.

0.8ml of spheroplast suspension was transferred to a Corning, 50ml orange-capped tube, and the spheroplasts were pelleted by centrifuging at 2,100 rpm in a MSE-Chilspin centrifuge, for 5 minutes at 4°C, and the resultant supernatant was discarded. Simultaneously, 2 x 10^7 myeloma cells were washed in 10ml of DMEM, and then resuspended in 10ml of a DMEM solution containing 7% (w/v) sucrose and 1.5mM EGTA. The myeloma cells from this suspension were pelleted over the spheroplast pellet, and the supernatant was discarded.

The fusion of spheroplasts and myeloma cells was carried out in a tube which was standing in a 41°C water-bath. 0.6ml of a prewarmed (37°C) solution of 50% (w/v) PEG 1500 in DMEM was added slowly (over 60 seconds) to the myeloma-cell/spheroplast pellet, with gentle stirring. The fusion reaction was allowed to proceed by incubating for a further 60 seconds, before being diluted out by the gradual addition of a DMEM solution containing 1.8mM EGTA; 2ml was added over the first two minutes, followed by 8ml over the next three minutes, and then 10ml dropwise. The number of viable myeloma-cells was then counted. Usually, the fusion of spheroplasts from 0.8ml of spheroplast suspension (≈ 8 x 10^8 bacterial-cell equivalents) with 2 x 10^7 myeloma cells resulted in a myeloma-cell viability of 25 - 50%. The transfected myeloma-cells were pelleted, and then resuspended in recovery medium (DMEM containing 10% (v/v) HS, 10% (v/v) FCS, penicillin / streptomycin and kanamycin (100µg/ml)) at a cell density of 3 x 10^5 viable cells/ml. 1ml aliquots of this cell suspension were then plated out into wells of a 24-well plate. The cells were allowed to recover for 48 hours and then selection for stable transfectants was applied. This involved selecting for expression from the transfected *gpt*-gene (see sections 1.7. and 3.1.) ; ½ml of the spent cell-medium was aspirated, and 1ml of recovery medium, supplemented with XHMPA (250µg/ml xanthine, 15µg/ml hypoxanthine, 6µg/ml mycophenolic acid), was added.

2.3.2. Electroporation.

Electroporation was used as an alternative transfection-method. The procedure used was based on that described by Potter *et al.* (1984). Plasmids were isolated on CsCl density-gradients, and then linearised in the pBR322-derived sequence with an appropriate restriction-enzyme. 2 x 10^7 myeloma cells were washed in PBS (Ca^{2+}, Mg^{2+}), and then resuspended in 0.5ml of the same buffer containing the digested plasmids (0.12 mg/ml).

The DNA / cell mixture was pipetted into an electroporation chamber. This consisted of a plastic 1ml-cuvette, with strips of aluminium foil placed on the inner-faces of two opposite sides serving as electrodes. A high-voltage shock (3,000 volts, 300 watts, 300mA) was passed through the solution, by bringing the two aluminium-foil electrodes into contact with the power supply, which was provided by a Bio-Rad 3000/300 power-pack. The power supply was quickly broken-off by a cut-out mechanism in the power pack; no attempt was made to measure the duration of the voltage pulse. The cells were allowed to recover for 10 minutes on ice, and then the cell viability was determined. The cells were then resuspended in recovery medium and plated out, as described in section 2.3.1.

2.3.3. Determination of stable gpt$^+$-transfection frequency.

As described in section 2.3.1., selection for expression from the *gpt* gene was applied 48 hours after the transfection. The selective medium (XHMPA-supplemented recovery medium) was changed when the cells began to exhaust the medium. This was indicated by a colour change from red to yellow, as the medium became more acidic. Initially, particularly following transfection with plasmids which gave efficient expression from *gpt*, the medium was changed every 2 or 3 days. Then the selection process took effect and untransfected cells perished and, therefore, less-frequent feeding was required. When feeding, medium was added dropwise and carefully, so as to minimise the possibility of dispersing the colonies. Visible colonies appeared from about 10 days following the transfection experiment. The stable *gpt*$^+$-transfection frequency was recorded by counting the number of colonies visible to the naked eye, after 21 days.

2.3.4. Expansion of transfectant clones.

For analysis of monoclonal populations, cells were cloned (see section 2.2.3.) or, alternatively, isolated colonies in the wells were identified microscopically and picked using a Pasteur-pipette. In the latter case, clones were transferred in approximately 0.2ml of medium to individual wells of a 96-well microtitre plate. Clones were then expanded in wells of 24-well plates prior to routine culture.

Polyclonal populations were made, by dispersing the colonies produced from a transfection in the individual wells in which they were cloned, and then by pooling these well populations in 25cm² flasks.

All stable *gpt*⁺-transfectants were maintintained XHMPA-supplemented maintenance medium, except where stated otherwise.

2.4. Preparation of materials for use in immune-precipitations.

2.4.1. Raising antisera.

Antiserum against chicken-egg lysozyme was raised in a New-Zealand White Rabbit. The animal was inoculated subcutaneously at four sites (neck, back and haunches) with 100µg of the lysozyme, which was dissolved in 100µl of a cPBS solution containing 50% (v/v) Freund's complete adjuvant. After six weeks, a booster-injection of 50µg of the lysozyme in 100µl of cPBS (without adjuvant) was given. Six days after boosting, the rabbit was bled by the main-ear artery using a 21G sterile hypodermic-needle. The blood was allowed to clot at 4°C for three hours, and then centrifuged at 4°C in a MSE-Chilspin bench centrifuge, at 4,000 rpm for 10 minutes. The serum (the supernatant) was removed, and then tested for the presence of anti-lysozyme antibody in an Ouchterlony immunodiffusion / immune-precipitation test. 4.5ml of a molten 1% (w/v) agarose solution in NET-buffer was pipetted onto one side of a glass microscope-slide, and allowed to solidify. Wells were cut in the agarose using a seven-hole punch. A 10µl aliquot of serum was pipetted into the centre well and serial dilutions (ranging from 1mg/ml to 0.1µg/ml) of lysozyme (the antigen) solution were pipetted into the outer wells. The immunodiffusion was allowed to proceed overnight, at 37°C in a humidified

atmosphere. When an immune-precipitate was produced (see section 4.7.4.), it was visible as a narrow white band between the antibody and antigen containing wells. Following confirmation that the serum was immune, a large-scale bleed was taken, and the antiserum was isolated and then aliquoted for storage at -20°C.

2.4.2. Preparation of fixed Staphylococcus aureus.

A colony of *Staphylococcus aureus* (Cowan strain-serotype 1 NCTC 8530), growing on a LB-agar plate (see section 2.6.1.), was picked and added to 25ml of LB-broth (see section 2.6.1.). This starter culture was grown for 28 hours at 37°C, and then 10ml of this culture was added to a further 490ml of LB-broth. This subculture was grown for 24 hours, and then the bacteria were harvested by centrifugation at 10,000 rpm, for 10 minutes at 4°C, in a MSE (6L) low-speed centrifuge. The bacterial pellet was resuspended in a 10% (v/v) formaldehyde / 0.2% (w/v) azide solution in PBS, and incubated on an orbital shaker for 90 minutes at room-temperature. Any remaining viable cells of this pathogenic bacterium were then killed, by heating to 80°C for 5 minutes and then cooling on ice. The bacterial-cell density was adjusted to 10% (v/v) in PBS / 0.2% (w/v) azide solution, split into 1ml aliquots and stored at -70°C. Before use in immune-precipitation reactions, cells in the 1ml aliquots were washed twice with 1ml of RIPA-buffer (150mM NaCl, 50mM Tris.HCl (pH7.4), 5mM EDTA, 0.1% (w/v) BSA, 0.1% (w/v) SDS, 1.0% (v/v) NP40, 1.0% (w/v) Na deoxycholate), then once with 1ml of NET / BSA / NP40 solution (150mM NaCl, 50mM Tris.HCl (pH7.4), 5mM EDTA, 0.1% (w/v) BSA, 0.05% (v/v) NP40), and then resuspended in the NET / BSA / NP40 solution at 10% (v/v).

2.5. Detection of proteins encoded by transfected genes.

2.5.1. Biosynthetic-radiolabelling of cell protein.

Between 3 x 10^5 and 5 x 10^5 cells were washed with 0.5ml of methionine-free Eagle's medium (L-glutamine- and L-methionine-free Eagle's modified minimum-essential medium, containing 2.0g/l sodium bicarbonate). The cells were then resuspended in 200µl of methionine-free Eagle's medium,

containing 2.5% (v/v) HS, 2.5% (v/v) FCS, 2mM glutamine, and 0.037-0.74 MBq (1-20 µCi) of ^{35}S-methionine (from a 50TBq/mmol, 555Mbq/ml stock-solution). They were then incubated in microtitre-plate wells for between 4 and 8 hours. The cells were resuspended from the bottom of the plate and the resultant cell-suspension was centrifuged. The cell pellet and supernatant were separated, and frozen at -20°C until subsequent analysis.

In order to obtain an absolute measurement of protein secretion-levels (see section 4.7.6. and table 4.2.), biosynthetic-radiolabelling was performed using defined medium which does not require serum supplementation; 0.74 MBq of ^{35}S-methionine was added to 200µl of Iscove's medium (with transferrin, albumin, soybean lipid, glutamine, 25mM Hepes and 3.024 g/l Na_2CO_3), and this medium was used to label the cells as described above.

2.5.2. Immune-precipitation.

Immune-precipitations of the radiolabelled proteins were carried out using a procedure based on that described by Kessler (1975). It made use of fixed *S. aureus*. The cell membrane of this bacterium contains protein-A which can act as an immunosorbant for rabbit Ig.

Immune-precipitations were performed directly on cell supernatants (i.e. secreted material), but for intracellular analysis, the cells were first lysed. Cell pellets were washed twice with 0.5ml of DMEM and resuspended in 100µl of DMEM. Aliquots were diluted 1:1 (v/v) with a lysis buffer (0.2M NaCl, 2mM $MgCl_2$, 0.1M Tris.HCl (pH7.4), 2% (w/v) NP40). The intracellular material was released by vortexing vigorously, and the insoluble material was then removed by centrifuging in an Eppendorf-5414 microcentrifuge, at 12,000 rpm for 5 minutes. The resulting supernatant and the cell supernatant were then immune-precipitated using the same procedure, as described below.

Usually, 5µl of an appropriate rabbit antibody was added to 50µl of radiolabelled-protein solution, and this was allowed to stand on ice for 60 minutes. Then 10µl of the fixed *S. aureus* suspension (see section 2.4.2.) was added, and the mixture was left on ice for a further 30 minutes. When different volumes of supernatants or lysates were used, the amounts of antibody and *S. aureus* were altered in proportion. The mixture

was then diluted to 1ml with NET / BSA / NP40 solution and the *S.aureus* was pelleted by centrifuging at 12,000 rpm for 1 minute. The pellet was then washed twice more in 1ml volumes of RIPA-buffer. Finally, the antibody and antigen were released from the *S. aureus* pellet, by boiling for three minutes in 2% (w/v) SDS solution (if the sample was to be analysed by scintillation counting), or by boiling in a solution containing 6M urea, 4% (w/v) SDS, 0.5M β-mercaptoethanol (if the sample was to be analysed by SDS-PAGE).

2.5.3. SDS-polyacrylamide-gel electrophoresis (SDS-PAGE).

Samples were prepared for SDS-PAGE by adding an equal volume of 2x sample-buffer (0.625M Tris.HCl (pH6.8), 2% (w/v) SDS, 0.7M β-mercaptoethanol, 5% (v/v) glycerol, 0.06% (w/v) bromophenol blue), and then boiling for 5 minutes. A standard protein molecular-weight marker track was run alongside the samples on the gels. This contained β-galactosidase (116KD), BSA (66KD), ovalbumin (45KD), carbonic anhydrase (29KD) and α-lactoglobulin (14.2KD).

SDS-PAGE was carried out using non-gradient slab gels, following a modification of the method described by Laemmli (1970). The gel system consisted of a stacking-gel containing 4% (w/v) acrylamide, and a resolving-gel containing 8, 11 or 13% (w/v) acrylamide. The gels were made by mixing the following solutions.

	Resolving-gel	Stacking-gel
30% (w/v) acrylamide / 2% (w/v) bis., (deionised)	8ml, 11ml or 13ml	4ml
1.5M Tris.HCl (pH8.8)	7.5ml	
0.5M Tris.HCl (pH6.8)		5ml
10% (w/v) SDS	0.3ml	0.2ml
20% (w/v) ammonium persulphate (fresh)	150µl	100µl
TEMED	20µl	13µl
H_2O	to 30ml	to 20ml

2.6. Bacterial cultures.

2.6.1. Growth medium.

All water used in media for growing bacteria had been distilled and passed through a Milli-Q water-purification system. All bacterial cultures were grown at 37°C with agitation, except for *Micrococcus lysodeikticus*, which was grown at 30°C. The liquid and solid media described below were used for growing bacteria.

Luria-Bertani medium (LB) (pH7.0) was usually used to grow *E. coli* cells. LB-broth contained 10g of bacto-tryptone, 5g of bacto-yeast extract and 5g of NaCl, per litre of water. LB-agar plates were prepared by adding bacto-agar (10g/l) to LB-broth, dissolving the agar by autoclaving, and then pouring the molten solution into petri-dishes before it solidified. For soft overlays (used when plating out pUC and M13 recombinants, as described in sections 2.8.2. and 2.9.4. respectively), 6g/l bacto-agar was used. Antibiotic-resistance imposed by transformed plasmids was selected for by adding ampicillin (50μg/ml) or tetracycline (15μg/ml), as appropriate, to the growth medium.

Minimal medium was made by mixing the following pre-sterilised solutions:

10x M9 medium (60g/l Na_2HPO_4, 30g/l KH_2PO_4, 5g/l NaCl, 10g/l NH_4Cl)	10ml
1M $MgSO_4$	0.1ml
20% (w/v) glucose	1.0ml
1% (w/v) thiamine (vitamin B1)	0.1ml
0.1M $CaCl_2$	0.1ml
H2O	to 100ml

Minimal-medium agar plates, containing 2% (w/v) bacto-agar, were made by adding molten agar-solution to a mixture of the pre-sterilised solutions.

SOB medium (pH 6.8 - 7.0), which was used in the preparation of transformation-competent cells, contained 2% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$ and 10mM $MgSO_4$.

2.6.2. Bacterial strains.

The following strains of *E.coli* were used as hosts for amplification of plasmid DNA:

Strain	Genotype	Reference
TG1	K12, (lac-pro)$^-$, supE, thi,hsdD5/F', traD36, proA$^+$B$^+$, lacI^q, lacZ, M15$^-$.	Gibson (1984)
DH5	F$^-$, endA1, hsdR179(rk$^-$, mk$^+$, supE44, thi-1, λ^-, recA1, gyrA96, relA1.	Hanahan (1985)

2.6.3. Long-term storage of bacteria.

Many *E. coli* transformants, harbouring important plasmids, were generated throughout the course of the project, and these were preserved by cold storage in glycerol. Following identification of the required recombinant-plasmid, the bacterial culture harbouring the plasmid was plated-out on a LB-agar plate containing the appropriate antibiotic. A single colony was picked from this plate, and then grown overnight in LB-broth / antibiotic. The culture was centrifuged in a MSE-Chilspin centrifuge, at 4,000 rpm for 5 minutes. The pellet was resuspended in 1ml of 10mM $MgSO_4$ and recentrifuged. The pellet was then taken up in 1ml of a solution containing 10mM $MgSO_4$ and 50% (v/v) glycerol, and stored at -20°C. The cells in such glycerinated suspensions remained viable throughout the investigation.

2.7. DNA manipulation.

2.7.1. General handling of DNA solutions.

To prevent nuclease-contamination of DNA solutions, all solutions were made up with Milli-Q purified distilled-water and then autoclaved. Most DNA-manipulations were carried out using disposable pipette-tips and 1.5ml microfuge-tubes, which were autoclaved prior to use. Sterile rubber-gloves were worn and aseptic technique was used throughout.

2.7.2. Isolation of plasmid DNA.

2.7.2.1. Small-scale isolation of plasmid DNA (mini-preps).

The method used was basically the alkaline-lysis method of Ish-Horowicz (1982). 10ml of LB-broth was inoculated with a single colony from an agar plate, or with a bacterial glycerol-suspension. The culture was then grown for a period of between 6 and 16 hours. 1.5ml of the culture was transferred to a 1.5ml microfuge-tube, and then centrifuged in an Eppendorf-5414 microcentrifuge, for 1 minute at 12,000 rpm. The pellet was resuspended in 100µl of ice-cold GET / lysozyme solution (50mM glucose, 10mM EDTA, 25mM Tris.HCl (pH8.0) and freshly dissolved lysozyme (4mg/ml)) and left at room temperature for 5 minutes. 200µl of 0.2M NaOH / 1% (w/v) SDS solution was then added and mixed by inverting the tube. After leaving the tube on ice for 5 minutes, 150µl of an ice-cold K acetate solution (≈pH4.8) (made by mixing 60ml of 5M K acetate, 11.5ml of glacial acetic-acid and 28.5ml of H_2O) was added. The tube was vortexed gently in an inverted position, and then placed on ice for 5 minutes. A white precipitated-complex of protein, cell debris and high molecular-weight DNA formed, and this was pelleted by centrifuging for 5 minutes in the microcentrifuge. 400µl of the supernatant was carefully removed, avoiding precipitated material, and transferred to a fresh tube. Phenol-extraction (see section 2.7.3.) was then necessary to remove residual protein. The meticulousness with which this step was performed depended on the intended use for the DNA being prepared. If a series of recombinants were simply to be screened by analysing restriction-enzyme digest patterns (see section 2.7.6.), a single extraction with an equal volume of phenol / [chloroform/iso-amyl-alcohol (24:1 (v/v))] (1:1 (v/v)) was sufficient. However, if the DNA was to be used for subcloning manipulations, two phenol-extractions, followed by a chloroform/iso-amyl-alcohol(24:1 (v/v))-

extraction, were performed. 300µl of the upper aqueous-phase was recovered following the extraction(s), and nucleic-acids were precipitated, by adding two volumes of ethanol, vortexing, and standing at room-temperature for 5 minutes. The resultant DNA / RNA precipitate was pelleted by centrifuging for two minutes in the microcentrifuge. The white pellet was washed with 1ml of 70% (v/v) ethanol, and then dried for 15 minutes in a vacuum desiccator. The pellet was dissolved in 50µl of TE-buffer (10mM Tris.HCl (pH8.0), 1mM EDTA) containing DNase-free pancreatic RNase (20µg/ml). The resultant DNA-solution was used directly in analysis or manipulation.

2.7.2.2. Large-scale isolation of plasmid DNA.

Bacterial cells containing the plasmid of interest were grown as an overnight culture in 10ml of LB-broth (with antibiotic). 0.5ml of this culture was added to 49.5ml of fresh-LB broth (with antibiotic) and grown for approximately 2½ hours, until the OD_{650nm} read 0.6-0.8. A 100mg/ml stock-solution of chloramphenicol, dissolved in ethanol, was added, to give a final chloramphenicol concentration of 0.2 mg/ml. The plasmid copy-number was then allowed to amplify by incubating for a further 16 hours.

The cells were harvested by centrifuging in a 50ml, Corning orange-capped tube at 4,000 rpm, for 5 minutes at 4°C. They were then washed in 5ml of buffer-A (50mM NaCl, 50mM Tris.HCl (pH8.0), 5mM EDTA) and repelleted. Then the cells were resuspended in 1ml of buffer-A containing 25% (w/v) sucrose. Then 200µl each, of ethidium bromide (EtBr) solution (5mg/ml) and freshly-made lysozyme solution (5mg/ml) were added, and mixed, before standing the tube on ice. After 5 minutes, 0.4ml of 0.25M EDTA (pH8.0) was added and mixed gently but thoroughly, and the tube was then left on ice for a further 5 minutes. Then 1.8ml of a lysis buffer (1% (w/v) Brij 35, 0.4% (w/v) Na deoxycholate, 62.5mM EDTA and 50mM Tris.HCl (pH8.0)) was added, mixed gently, and the the tube was stood on ice for 30 minutes with intermittent shaking. The mixture was then centrifuged in a Sorvall-SS34 rotor at 20,000 rpm, at 4°C for 30 minutes. This spin removed the chromosomal DNA and bacterial debris, which formed a tight pellet on the bottom of the tube. As much of the supernatant was removed as possible and then made up to a volume of 4.25ml by addition of TE-buffer (pH8.0). 4.35g of ultrapure CsCl was added and dissolved by gentle agitation, and then

227µl of EtBr solution (10mg/ml) was added. Carefully avoiding EtBr / protein aggregates which formed, the solution was transferred into a Beckman quick-seal 13 x 51mm polyallomer-tube. The tube was sealed, and then centrifuged in a VTi65.2 vertical-rotor at 50,000 rpm, for 16 hours at 18°C, using a Beckman (L5-65) ultracentrifuge. During the spin, a CsCl density-gradient formed and this resulted in DNA collecting in two bands; the upper band consisted of linear bacterial-DNA and nicked-circular plasmid DNA; the lower band consisted of closed-circular plasmid DNA. An air-hole was made at the top of the tube with a pair of scissors, and then a hypodermic-needle (19 x 1½ mm) was inserted through the wall of the tube just below the lower band. The band was slowly removed using a syringe. The DNA solution was transferred to a 1.5ml microfuge-tube, and the EtBr was removed by successive extractions with equal volumes of water-saturated butan-1-ol, until the pink colour was no longer visible. Finally, the aqueous phase was dialysed against several changes of TE-buffer (pH8.0), and then the solution was stored at -20°C.

2.7.3. Phenol- and chloroform-extractions.

Liquefied phenol was supplied by Fisons and was sufficiently pure as for redistillation to be unnecessary. When phenol was received, the antioxidant, 8-hydroxyquinoline, was added to a final concentration of 0.1% (w/v), and the resultant solution was stored in 100ml aliquots at -20°C. When required, phenol was thawed out. It was then neutralised by extracting twice with an equal volume of 1M Tris.HCl (pH 8.0). The pH of the aqueous phase was checked to be greater than 7.6, and then the Tris was diluted by equilibriating the phenol layer three times with an equal volume of TE-buffer (pH 7.4). This equilibriated phenol was stored at 4°C in the dark, for up to 3 months.

Nucleic-acid solutions were rendered protein-free (e.g. after Klenow-extensions and restriction digests, and during mini-preps) by extraction with an equal volume of equilibriated phenol. Extractions were usually carried out in 1.5ml microfuge-tubes. The tube was vortexed for ten seconds so that an emulsion between the phenol and aqueous phases was formed, and then the phases were separated by centrifuging, at 12,000 rpm for 3 minutes, in an Eppendorf microcentrifuge. However, phenol-extractions

on larger aqueous-samples were carried out in glass universal-bottles, and the phases were separated by centrifuging, for at least 15 minutes at 3,000 rpm, in a MSE bench-centrifuge. The upper aqueous-phase, which contained the DNA, was removed, carefully avoiding carry-over of denatured protein from the phase interface. When handling small quantities of DNA, recovery was enhanced by back-extraction following removal of the aqueous phase. An equal volume of fresh TE-buffer (pH 8.0) was added to the phenol phase and the phases were mixed and separated as above. Then the resultant second aqueous-phase was combined with the first. When highly pure nucleic-acid preparations were required, phenol-extractions were repeated until there was no visible protein at the interface of the phenol and aqueous phases. Following phenol-extraction(s), an equal volume of chloroform/iso-amyl alcohol (24:1 (v/v)) was added to the aqueous phase to remove residual phenol. The mixture was vortexed and then the phases were separated by centrifuging (as for phenol-extractions). The aqueous phase, which contained the nucleic-acid, again formed the upper layer.

2.7.4. Ethanol-precipitation of nucleic-acids.

DNA was recovered from aqueous solution by ethanol-precipitation. Usually, tRNA-carrier was added to the DNA solution to give a concentration of 30μg/ml. This was followed by addition of 8M ammonium-acetate solution to give a final concentration of 2.7M. Then two volumes of ethanol were added to this DNA solution. The solution was mixed well, and then chilled in a dry-ice / isopropanol bath for thirty minutes. The DNA precipitate which formed was collected by centrifuging, for five minutes at 12,000 rpm, in an Eppendorf microcentrifuge. The DNA pellet was dried in a vacuum desiccator, which was attached to a Cole-Palmer aspirator pump. Then the DNA was resuspended in an appropriate volume of TE-buffer (pH8.0).

Single-stranded DNA (synthetic oligonucleotides and M13) and RNA were found to precipitate more efficiently when 2.5 volumes of ethanol were added.

2.7.5. Quantitation of DNA and RNA.

The concentration of nucleic-acid solutions was determined spectrophotometrically, by measuring the absorbance at a wavelength of 260nm. The concentration was calculated using the information that an OD_{260nm} of 1.0 corresponds to approximately 50µg/ml of double-stranded DNA, 40µg/ml of single-stranded DNA and RNA, and 20µg/ml of oligonucleotides. The purity of nucleic-acid solutions was also estimated, by measuring the absorbance at 280nm and calculating the OD_{260nm}/OD_{280nm} value. Pure preparations of DNA and RNA have OD_{260nm}/OD_{280nm} values of 1.8 and 2.0 respectively. If the values obtained differed significantly from these expected values, further phenol-extractions were performed in order to remove residual protein.

2.7.6. Restriction-endonuclease digestion of DNA.

Restriction-endonuclease digestions of plasmid, M13 RF and genomic DNA, were carried out using the buffer and temperature conditions recommended by the supplier of the enzymes, BRL. Restriction-endonuclease buffers were made as 10x stock-solutions which were stored in aliquots at -20°C. The buffers and their compatible enzymes are listed below:

Restriction Buffer	Restriction Endonuclease
Buffer I - Low salt (10x)	
0.5M Tris.HCl (pH 8.0) 0.1M $MgCl_2$	*Acc* I, *Cla* I, *Rsa* I.
Buffer II - Medium salt (10x)	
0.5M Tris.HCl (pH 8.0) 0.1M $MgCl_2$ 0.5M NaCl	*Hind* III, *Mbo* I, *Nde* I, *Pst* I, *Pvu* II, *Sst* I, *Xba* I, *Xho*
Buffer III - High salt (10x)	

0.5M Tris.HCl (pH 8.0) 0.1M $MgCl_2$ 1M NaCl	*Bam* HI, *Bgl* II, *Eco* RI, *Nco* I, *Sal* I.

Buffer IV - Sma I (10x)

0.2M Tris.HCl (pH 7.4) 50mM $MgCl_2$ 0.5M KCl	*Apa* I, *Hinc* II, *Kpn* I, *Sma* I.

BSA was also added to a final concentration of 100µg/ml in the digestions. Digestions were carried out in 0.5ml or 1.5ml microfuge-tubes, and for between 1 and 16 hours depending on the amount of DNA to be digested. The digests were performed in a minimum volume, which was usually between 10 and 50µl. The volume was determined by the requirement of keeping the glycerol concentration below 5% (v/v), because higher concentrations can inhibit the enzyme; the enzymes were supplied in 50% (v/v) glycerol (which allows them to be stored, unfrozen at -20°C) and, therefore, the amount of enzyme solution added as a proportion of total volume never exceeded 10% (v/v). Mini-prep prepared DNA did not digest as efficiently as DNA isolated from CsCl density-gradients. Therefore, when digesting mini-prep prepared DNA, ten times the recommended amount of enzyme was used. Multiple digests were carried out simultaneously in the same buffer whenever this was possible. However, if the enzymes had radically different optimum-buffers, a two-stage digestion was necessary. The enzyme favouring the lower-salt buffer was added first, allowed to digest, and then the buffer concentration was appropriately modified prior to addition of the second enzyme. Aliquots of the digests were analysed by agarose-gel electrophoresis to determine the sizes of the DNA fragments (see section 2.7.7.). Such analysis was performed to check the extent of digestion during subcloning manipulations, and also to screen recombinants by 'restriction-endonuclease mapping' to check if recombinants had the desired structure.

2.7.7. Agarose-gel electrophoresis.

2.7.7.1. Analytical gels.

The DNA fragments generated by restriction-endonuclease cleavage were separated, according to size, by electrophoresis in horizontal agarose-gels. Molten agarose (Miles) - solution (0.7-2.0% (w/v)) was prepared in E-buffer (40mM Tris.HCl (pH7.8), 30mM acetate, 1mM EDTA). E-buffer was made by mixing 9.7g of Tris. base, 0.75g of Na_2EDTA, 0.4g of NaOH and 36ml of glacial acetic-acid, and then making up to 1 litre with water. The agarose solution was poured onto perspex gel-formers, and loading wells were formed by insertion of plastic combs into the molten agarose-solution. The size of the gel depended on the number of samples and the required resolution; either minigels (8 x 8 cm) or midigels (11 x 14 cm) were used. Once solidified, the gels were transferred to a horizontal electrophoresis-tank and immersed in E-buffer. One-sixth volume of 6x loading-buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanole and 30% (v/v) glycerol) was added to DNA samples (usually 10-20µl), before loading into the wells. Gels were run at 80-100 volts for 1-2 hours, or at 8 volts, overnight. Following electrophoresis, the DNA fragments in the gels were stained by transferring the gel to an EtBr solution (0.5µg/ml) for twenty minutes. The banded fluorescence-pattern from DNA fragments with bound EtBr was visualised under UV-light (provided by a Fotodyne (New Berlin, Wisconsin) transilluminator). The band pattern was recorded by photography using Polaroid instant or Kodak T-max film, with a Polaroid MP-4 land-camera which was fitted with an orange filter. The sizes of DNA fragments were estimated by comparison of mobility with co-electrophorised *Hind* III-digested bacteriophage lambda (*c*I857S am7) DNA (fragment sizes: 23.1, 9.4, 6.6, 4.4, 2.3, 2.0, 0.56, Kb pairs), or *Hind* III / *Eco* RI digests of the same bacteriophage DNA (fragment sizes: 21.2, 5.2, 5.1, 4.3, 3.5, 2.0, 1.9, 1.6, 1.4, 1.0, 0.9, 0.56, Kb pairs).

2.7.7.2. Preparative gels for isolation of DNA fragments.

The preparation of samples, gel loading and electrophoresis, was as described for analytical gels, except Miles agarose was not used. Instead, Sigma type-II (medium EEO) agarose was used when purified DNA fragments were to be used in further DNA manipulation, and BRL ultra-pure, low-melting-point agarose was used when isolating radiolabelled probes. In the former case, DNA bands were visualised by the fluorescence of bound

EtBr under longwave UV-light, and in the latter case, the required fragment was located by autoradiography.

Many procedures were compared for recovery of DNA from preparative gels, which was intended for subsequent DNA-manipulation. The two procedures described below allowed the DNA manipulations to be carried out with the greatest efficiency.

2.7.7.2.1. Electroelution into dialysis bags.

The electroelution method was originally described by McDonnell *et al.* (1977). A small slice of agarose, containing the DNA fragment, was cut out of the gel with a scalpel, and then placed in a dialysis-bag containing approximately 200µl of E-buffer. The bag was immersed in a shallow layer of E-buffer contained in a horizontal electrophoresis-tank. An electric current (100 volts) was passed through the bag for 30 minutes. Then the DNA was released from the walls of the dialysis bag, by reversing the polarity of the current for 30 seconds. The buffer inside the bag was removed, carefully avoiding the gel slice, and the inside of the bag was then washed out with a further 50µl of E-buffer. The buffer was then passed through a small column of siliconised glass-wool, in a Pasteur-pipette, to remove any pieces of gel. Three phenol- and two chloroform/iso-amyl-alcohol-extractions (see section 2.7.3.) were performed on the buffer solution to remove sulphated polysaccharides, which are eluted from the agarose. Then the DNA was recovered by ethanol-precipitation. Electroelution was also used to recover small single-stranded DNA fragments which were separated on polyacrylamide gels (see section 2.10.1.3.). The same procedure was used, except the buffer was 0.5x TBE, and phenol- and chloroform/iso-amyl-alcohol-extractions were not necessary.

2.7.7.2.2. 'Geneclean'.

The use of the 'Geneclean' kit, supplied by Stratech Scientific Ltd., was found to be the most efficient method of recovering DNA from agarose-gel slices, which could then be easily manipulated in subcloning procedures. It was especially efficient for recovery of fragments smaller than 7Kb. The kit makes use of solutions of undeclared composition and/or

concentration. The process involves the binding of DNA to glass in the presence of NaI, followed by washing in a high-salt buffer, and then elution of the DNA in a low-salt buffer or water. It therefore uses many of the principles described by Vogelstein and Gillespie (1979). The procedure involved dissolving the gel slice in 2½ volumes of a NaI solution, by heating to 50°C. The solution was then cooled, 5µl of a 'glassmilk' suspension was added and then the tube was left on ice for 5 minutes. The resultant glassmilk / DNA complex was then pelleted by pulse-centrifuging for 10 seconds in an Eppendorf microcentrifuge, and washed three times in 200µl volumes of ice-cold 'new' solution. The 'glassmilk' was then resuspended in H_2O, and heated to 50°C to release the DNA.

2.7.8. Blunt-ending of overhangs generated by restriction-endonuclease digestion.

3'-recessed termini (5'-overhangs) were 'filled-in' using the large-fragment of DNA-polymerase I (Klenow). Klenow's 5'→3' polymerase-activity resulted in the recessed strand being extended as far as the end of the 5'-overhang. A typical reaction was carried out at 26°C for 30 minutes, in a reaction volume of 10µl, as illustrated below:

DNA	0.2µg
10x fill-in buffer (0.5M Tris.HCl (pH7.5), 0.1M $MgCl_2$, 4mM DTT, 10mM spermidine)	1µl
2.5mM dATP/dCTP/dGTP/dTTPs	1µl
BSA (1mg/ml)	0.5µl
Klenow	2 units
H_2O	to 10µl

If the DNA was to be further digested with another restriction-endonuclease, it was necessary to first inactivate / remove the Klenow, to prevent it also filling-in overhangs produced in these digestions. Klenow was heat inactivated, by adding EDTA (pH8.0) to a final concentration of 25mM and then heating to 65°C for thirty minutes. The volume of the solution was then increased to 100µl by adding TE-buffer (pH8.0), and three phenol-extractions and one chloroform/iso-amyl-alcohol - extraction were then performed. The DNA was then recovered by ethanol-precipitation. However, if a fill-in reaction was to be followed by a ligation reaction,

the removal of Klenow was unnecessary. ATP and T4 DNA-ligase were simply added to the reaction mix at the same concentrations as used in standard ligation-reactions (see section 2.7.11.).

3'-overhangs were 'cut-back' to generate blunt-ended DNA using one of two alternative methods. The 3'→5' exonuclease-activity of Klenow could be employed, using identical reaction conditions to those described above for filling-in 5'-overhangs. However, T4 DNA-polymerase was found to have a more efficient 3'→5' exonuclease activity. A typical reaction was carried out at 37°C for one hour, in a reaction volume of 10µl, as illustrated below:

DNA	0.2µg
10x T4 DNA-polymerase buffer (0.5M glycine (pH7.5), 60mM $MgCl_2$, 100mM DTT, 65µM EDTA, 160mM ammonium sulphate)	1µl
1µl 2.5mM dATP/dCTP/dGTP/dTTPs	1µl
BSA (1mg/ml)	1µl
T4 DNA-polymerase	5 units
H_2O	to 10µl

2.7.9. Nuclease-Bal31 digestion.

Bal31-exonuclease from *Alteromonas espejiana* acts on double-stranded DNA and degrades both strands at approximately the same rate (Legerski *et al.*, 1978). It was used in a situation when a deletion of a sequence was required, but no convenient restriction-sites were available. This involved the removal of the native translation-initiation codon (ATG) of an IgH-gene - promoter fragment (see section 3.3.1.). This codon was contained within a *Nco* I recognition site and was flanked downstream (3'-) by a series of recognition sites of the pUC8 polylinker-sequence. The plasmid was digested with *Nco* I and ethanol-precipitated. On resuspending, a Bal31 digest was performed, at 30°C, using the following reaction mixture, which was adapted from the procedure described by Legerski *et al.* (1978):

DNA	500ng
2x Bal31-buffer (0.4M NaCl, 40mM Tris.HCl (pH8.0), 24mM $CaCl_2$, 24mM $MgCl_2$, 2mM EDTA)	50µl
BSA (1mg/ml)	25µl

Bal31 exonuclease	2.0 units
H_2O	to 100µl

Aliquots were taken at time intervals during the digest, and the reaction was stopped, by the addition of EGTA to a final concentration of 20mM and standing on ice. Analysis by agarose-gel (1% (w/v)) electrophoresis did not detect any reduction in the size of the linearised plasmid following a 1 hour digestion with Bal31, indicating that the reaction was slow. Presumably, the low activity of the enzyme was due to impurities in the reaction mix, because the DNA had been prepared by the alkaline-lysis method (see section 2.7.2.1.). The Bal31 was removed by phenol- and chloroform/iso-amyl-alcohol - extractions, and then the DNA was ethanol-precipitated. Uneven DNA-ends were filled-in or cut-back with Klenow, and then ligated with T4 DNA-ligase. This DNA was used to transform TG1 cells (see section 2.8.). The extent of the plasmid deletions were determined by preparing plasmid DNA from transformants, and then restriction mapping (see section 2.7.7.1.). The only recombinant plasmids containing detectable deletions, were those derived from transformants which had been transformed with the DNA which had been Bal31 treated for at least 20 minutes. A recombinant was identified which had lost three polylinker recognition-sites downstream of the *Nco* I site (a deletion of at least 12 base pairs), but retained the most-3' polylinker recognition-sequence (*Hind* III). As Bal31 is reported to digest both ends of DNA simultaneously, an assumption was made that the ATG translation-initiation codon had also been deleted, because only a one base-pair deletion was required at the other end of the *Nco* I - digested plasmid to destroy this codon.

2.7.10. Dephosphorylation of restriction fragment 5'-termini.

Some subcloning manipulations required the insertion of a DNA fragment into a backbone, which itself, had two cohesive ends. Such ends were generated when a plasmid was cut with a single restriction-endonuclease, or when cutting with more than one restriction endonuclease which generated cohesive ends (e.g. *Bam* HI and *Bgl* II). During an *in-vitro* ligation reaction, the self-ligation of such backbones occurs preferentially over intermolecular-ligation with an insert fragment.

Therefore, to minimise the number of recombinants that had to be screened, self-ligation was minimised. This was done by removal of the 5'-terminal phosphate groups from the backbone DNA-fragment, using calf-intestinal alkaline phosphatase (CIP). This prevented self-ligation, because the 5'-phosphate group is essential for formation of the phosphate bridge with the 3'-OH group at the other DNA terminus. Therefore, the backbone fragments could only ligate with non-CIP-treated insert fragments. The reaction was carried out as described below.

DNA (≈100ng) was precipitated following restriction-digestion and then incubated for two hours, at 37°C with 20 units of CIP, in 10mM Tris.HCl (pH9.2), 0.1mM EDTA. The dephosphorylated DNA was ethanol-precipitated and then redissolved in TE-buffer (pH8.0). Plasmid DNA transforms (see section 2.8.) bacteria much more efficiently when it is circular than when it is linearised. Therefore, the efficiency of the dephosphorylation reaction was estimated, by attempting to self-ligate an aliquot of the DNA solution in a ligation reaction (see section 2.7.11.), and then determining the transformation frequency of DNA produced by the ligation reaction (see section 2.8.). If the frequency was low, relative to that produced by DNA treated identically except without CIP treatment, the remaining dephosphorylated DNA was mixed with the insert DNA in a ligation reaction.

2.7.11. Ligation of DNA fragments.

Linear fragments of DNA, containing terminal 5'-phosphate and 3'-OH groups, were ligated to produce circular recombinant-plasmids, which were then transformed into bacterial cells (see section 2.8.). Ligations were usually carried out overnight, at 14°C in a volume of 10µl, using the reaction mixture illustrated below:

DNA	50ng
5x ligation-buffer (0.25M Tris.HCl (pH 7.4), 50mM $MgCl_2$, 50mM DTT, 5mM spermidine)	2µl
10mM ATP	1µl
BSA (1mg/ml)	1µl
T4 DNA-ligase	1 unit
H_2O	to 10 µl

Ligations were performed, with gel-purified DNA fragments (see section 2.7.7.2.) or, alternatively, with fragments directly from restriction-endonuclease digestions ('shotgun cloning'), which were unpurified other than ethanol-precipitated (see below). In the latter situation, fragments additional to those whose ligation was desired were present and, consequently, many more unwanted recombinants were generated. Hence, more extensive screening of transformants by restriction-endonuclease mapping (see section 2.7.6.) was required to obtain the required recombinant.

Before two fragments were ligated, they were ethanol-precipitated and redissolved in TE-buffer (pH8.0). Then small aliquots of the DNA solutions were run on an agarose gel and the fragment concentrations were estimated by the intensity of EtBr fluorescence (see section 2.7.7.1.). Generally a 1:1 molar-ratio was used in the ligation, but an exception applied when inserting a fragment into a backbone with cohesive ends. Even when the backbones had been treated with CIP (see section 2.7.10.), the reactions had rarely gone to completion. Therefore, in order to minimise self-ligation, a ten-fold excess of the insert fragment over the backbone fragment was put into the ligation reaction, and higher DNA-concentrations (≈ 500ng/10µl) were used than usual. Similar conditions were also required when inserting fragments into blunt-ended backbones (e.g. pTK-E(a) construction, see section 3.3.1.).

2.8. Transformation of E.coli.

2.8.1. Preparation of competent cells.

Untreated bacterial-cells take up extracellular DNA (are transformed) very inefficiently. Therefore, cells were made highly receptive or 'competent' for efficient transformation using a method based on that described by Hanahan (1983).

A single colony of *E. coli* (strain TG1 or DH5) was picked off a minimal-medium agar plate, and used to inoculate 5ml of SOB medium. The culture was grown overnight, and then 1ml of the culture was subcultured into 99ml of fresh, warm SOB-medium, and grown for approximately two hours, until the OD_{650} read 0.5. The cells were pelleted by centrifuging in

Corning orange-capped tubes at 4,000 rpm, at 4°C for 5 minutes, in a MSE-Chilspin centrifuge. The pellets were resuspended and pooled in a 40ml total-volume of ice-cold transformation-buffer I (30mM K acetate, 100mM RbCl, 10mM $CaCl_2$, 50mM $MnCl_2$, 15% (v/v) glycerol, adjusted to pH5.8 with 0.2M glacial acetic-acid). The tubes were then left to stand on ice for five minutes. The cells were repelleted, and then resuspended in 4ml of ice-cold transformation-buffer II (10mM Mops., 75mM $CaCl_2$, 10mM RbCl, 15% (v/v) glycerol, adjusted to pH6.5 with 1M KOH). The cell suspension was transferred, in 100µl aliquots, into 1.5ml microfuge-tubes, and then flash-frozen on dry-ice. These competent cells were stored at -70°C, and thawed immediately before using in a transformation. The transformation efficiency of stored cells decreased with time. Therefore, fresh competent-cells were prepared when a high transformation-frequency was essential, such as following lengthy DNA-subcloning manipulations, where the amount of DNA was potentially limiting.

2.8.2. Transformation procedure.

Half the volume of a ligation-reaction mixture (5µl) was added to 100µl of competent-cell suspension (see above). The tube was left to stand on ice for 20 minutes. Then the cells were heat-shocked, by incubating the tube in a 42°C water-bath for two minutes. The cells were then plated out, but the procedure used depended on the backbone fragment(s) of the DNA used in the transformation. The paragraph below describes the procedure used for plasmids containing a pBR-derived backbone, which carried an antibiotic (ampicillin) - resistance gene, but the procedure used for M13 recombinants is described in section 2.9.4.

The transformed cells were first allowed to recover and express ampicillin-resistance, by incubating at 37°C for 45 minutes, in 1ml of LB-broth (without ampicillin). Then cells were usually plated out at two different concentrations (1% and 99% of cells), on LB-agar plates containing ampicillin (see section 2.6.1.). They were left to incubate at 37°C; ampicillin-resistant TG1 colonies became visible after about seven hours incubation. Following transformation with plasmids which contained a fragment inserted into the polylinker sequence of a pUC plasmid, a blue-white screen was used to distinguish recombinant plasmids

from the non-recombinant parent pUC-plasmid. Following the recovery period in LB-broth (see above), the cells were added to 2.5ml of molten LB-top-agar (see section 2.6.1.), which contained 50µl of X-gal (2% (w/v) in N,N'-dimethylformamide) and 10µl of 100mM IPTG, and which had been preincubated at 43°C. This agar solution was poured onto LB-agar plates (containing ampicillin), allowed to solidify, and then the plate was incubated at 37°C. Cells transformed with the parent pUC-plasmid gave rise to blue colonies, but cells transformed with recombinant plasmids gave rise to white colonies. The screen works by virtue of a β-galactosidase gene located on the pUC plasmid, whose expression is induced by IPTG. The expressed enzyme hydrolyses X-gal to a blue dye (bromochloroindole). Insertion of the foreign DNA into the polylinker sequence, which is situated within the β-galactosidase gene, interferes with production of β-galactosidase and, therefore, no hydrolysis of the dye occurs.

2.9. M13 cloning-vectors.

2.9.1. Introduction.

M13 is a rod-shaped bacteriophage which contains a circular single-stranded DNA (+ strand) genome. When the phage infects a bacterial cell, a double-stranded replication form (RF) of the DNA is synthesised. A rolling-circle mechanism of replication then amplifies the RF, which is used to produce further (+) strands which are ultimately packaged into phage particles. The phage particles are released from the cell by budding through the cell membrane. The M13 genome has been modified to produce M13 cloning-vectors, which can be carried through the same 'life cycle' as wild-type M13 DNA. Examples are the cloning vectors, M13mp9 (Vieira and Messing, 1982) and M13mp18 (Yanisch-Perron *et al.*, 1985). For simplification, M13 cloning-vector DNA will be referred to as M13 DNA and the phage carrying the M13 cloning-vector will be referred to as M13 phage. A polylinker sequence within the M13 DNA allows insertion of DNA fragments. The resultant M13-recombinant DNA can then be used for purposes such as DNA sequencing and preparation of primers and hybridisation probes (see section 2.10.)

A bacterial cell can only be infected by M13 phage when it has the F-pilus phenotype (genotype F'). Hence, all manipulations with M13 phage were performed using the *E. coli* host, strain TG1, which contains the F-pilus. TG1 was maintained, prior to M13 work, on minimal-medium agar plates (see section 2.6.1.) which select for maintenance of the F-pilus.

2.9.2. Growing M13-phage infected cultures.

M13 phage and derivative recombinants were stored at -20°C, in aliquots of supernatants derived from phage-infected bacterial cultures. 1µl of such a supernatant, or a plaque lifted from a top-agar plate (see section 2.9.4.), was added to 2ml of LB-broth. Then 20µl of an overnight culture of TG1 'feeder cells', which had been instigated by inoculating 10ml of LB-broth with a single colony picked from a minimal-medium agar plate, was added. Simultaneously, 80µl of a TG1 feeder-cell suspension was added to 8ml of LB-broth. These cultures were grown for 2 hours at 37°C, pooled, and then grown for a further 4 hours at 37°C. This culture was used to isolate double-stranded M13 DNA (see section 2.9.3.) or single-stranded M13 DNA (see section 2.9.5.).

2.9.3. Preparation of double-stranded (replicative form (RF)) M13 DNA.

As alluded to above, RF DNA is found inside the bacterial cells of a M13-phage-infected culture. 1.4ml of a phage-infected culture was transferred to a 1.5ml microfuge-tube, and the cells were pelleted by centrifuging, at 12,000 rpm for 1 minute, in a microcentrifuge. The supernatant was removed, and the DNA was isolated from the cell pellet using the alkaline-lysis method (see section 2.7.2.1.). *In vitro* manipulations of RF DNA, such as subcloning of DNA fragments into a M13 polylinker-sequence, were carried out using the same methods as for plasmid DNA.

2.9.4. Screening for M13 recombinants.

The M13 recombinants used in this study are described in section 2.10.1. They were isolated using the procedure described below. TG1 cells

were transformed with the products of a ligation reaction, in which a fragment had been inserted into the polylinker sequence of M13mp9 or M13mp18 RF-DNA. Following transformation (after the 42°C heat-shock, see section 2.8.2.), the cells were added to 2.5ml of molten LB-top-agar, which contained 50µl of X-gal (2% (w/v) in N,N'-dimethylformamide) and 10µl of 100mM IPTG, and which had been preincubated at 43°C. Then 50µl of a suspension of TG1 feeder-cells (see section 2.9.2.) was added. This top-agar was poured onto a LB-agar plate (without antibiotic), allowed to solidify, and then incubated at 37°C. The feeder-cells gave rise to a bacterial lawn, which contained plaques (areas of retarded bacterial growth) where infection with M13 phage had occurred. As in pUC plasmids (see section 2.8.2.), M13 DNA (the cloning vectors) contains a β-galactosidase gene which is disfigured by insertion of a fragment into the polylinker sequence. Therefore, the blue-white screen was used to allow identification of recombinant-M13 phages; plaques containing recombinant-M13 phage were colourless ('white') and plaques containing wild-type M13 phage were blue.

Recombinant phages were picked, by touching the surface of the plaque with a capillary tube, and transferred to a suspension of TG1 feeder-cells. The cultures were expanded (see section 2.9.2.), and then RF DNA was prepared from the infected cells (see section 2.9.3.). The recombinants were analysed by restriction-endonuclease mapping (see section 2.7.6.) to determine if they had the required structure.

2.9.5. Preparation of single-stranded M13 DNA.

As discussed above, in a M13-phage-infected culture, phage particles containing single-stranded DNA are released from the bacterial cells into the extracellular medium. The single-stranded DNA was purified as described below, carrying out all steps at room-temperature. Firstly, 1.5ml of a phage-infected culture was centrifuged in a microcentrifuge, at 12,000 rpm for 5 minutes, and 1ml of the supernatant was carefully removed, taking care to avoid the cell pellet. To ensure all cells had been removed, the supernatant was recentrifuged, and 800µl of supernatant was transferred to a fresh tube. Then 200µl of a 10% (w/v) PEG 6000 / 2.5M NaCl solution was added to the supernatant, and mixed in thoroughly. The phage particles

were allowed to precipitate by standing for thirty minutes, and then pelleted by centrifuging, at 12,000 rpm for 5 minutes, in a microcentrifuge. The supernatant was removed using a drawn-out Pasteur-pipette. The tube containing the pellet was then recentrifuged, and residual supernatant was removed as completely as possible. 100µl of TE-buffer (pH8.0) was added to the tube, and the phage pellet was resuspended by vortexing. Protein was then removed from the phage particles by a phenol-extraction. An equal volume of phenol was added, and the mixture was vortexed for 15 seconds. The mixture was allowed to stand for 15 minutes, and then revortexed. Then the phases were separated by centrifuging, at 12,000 rpm for 5 minutes, in a microcentrifuge. 90µl of the aqueous phase was removed and then added to an equal volume of chloroform/iso-amyl-alcohol. This mixture was vortexed, and then centrifuged, for 1 minute at 12,000 rpm, in a microcentrifuge. The aqueous phase was removed, and the single-stranded DNA was recovered by ethanol-precipitation using 2.5 volumes of ethanol. The small pellet was washed with 70% (v/v) ethanol, vacuum dried, and then resuspended in 50µl of TE-buffer (pH8.0).

2.10. Use of recombinant M13 single-stranded DNA.

Several M13 recombinants were constructed during this study which were valuable tools in the study of gene expression. Their construction and uses are described in the following sections.

2.10.1. Preparation of primers and hybridisation probes.

2.10.1.1. Construction of M13 recombinants.

DNA fragments, containing partial or complete coding sequences of the genes which were to be expressed in J558L, were subcloned into the polylinker sequence of either M13mp18 or M13mp9. In some cases it was a requirement that the probe or primer, to be produced using the single-stranded M13-recombinant template (see section 2.10.1.3.) be complementary to the RNA encoded by the gene. Therefore, recombinants were generated and identified, in which the insert was cloned in the orientation which would allow the RNA-complementary single-stranded DNA probe to be generated. M13

recombinants were constructed as described below using plasmids obtained from sources listed in section 2.1.2., and are illustrated in figure 2.1.

M13mp18-ova.

A 1.4Kb *Hind* III ovalbumin-cDNA fragment (from pTK_2ov) was cloned into the *Hind* III site of the M13mp18 polylinker-sequence. A recombinant, with the required orientation, was identified by restriction-endonuclease mapping, making use of a *Pvu* II site situated approximately 0.5Kb from the 5'-end of the cDNA fragment. The single-stranded probe generated using this recombinant (see section 2.10.1.3.) was used to analyse for genome-integrated ovalbumin sequences, and ovalbumin transcripts (see sections 4.2.3., 4.2.4., 4.7.2. and 4.7.3.)

M13mp18-lys.

A 0.5Kb *Hind* III chicken-lysozyme-cDNA fragment (from pTK_2lys) was cloned into the *Hind* III site of the M13mp18 polylinker-sequence. A recombinant, with the required orientation, was identified by DNA-sequencing. The single-stranded probe generated using this recombinant (see section 2.10.1.3.) was used to analyse for lysozyme transcripts (see section 4.7.3.).

M13mp9-gpt.

A 1.9Kb *Hind* III - *Bam* HI fragment (from pSV_2gpt), which contains the *gpt* coding-sequence and downstream SV40 early-region sequences, was cloned into the *Hind* III and *Bam* HI sites of the M13mp9 polylinker. The single-stranded probe generated using this recombinant (see section 2.10.1.3.) was used to analyse for genome-integrated *gpt*-SV40 sequences (see sections 4.2.3., 4.2.4. and 4.7.2.).

M13mp9-5'gpt.

A 0.12Kb *Hind* III - *Bgl* II (filled-in) fragment from pSV_2gpt, which was derived from the 5'-end of the *gpt* gene, was cloned into the *Hind* III and *Sal* I (filled-in) sites of the M13mp9 polylinker. The recombinant containing the insert was easily identified by restriction-endonuclease digestion, due to the regeneration of the *Bgl* II site from the ligation of filled-in *Sal* I and *Bgl* II ends. The single-stranded sequence, generated using this recombinant (see section 2.10.1.3.), was used as a primer in primer-extension analysis of *gpt* transcripts (see section 3.6.).

2.10.1.2. Oligonucleotide primers.

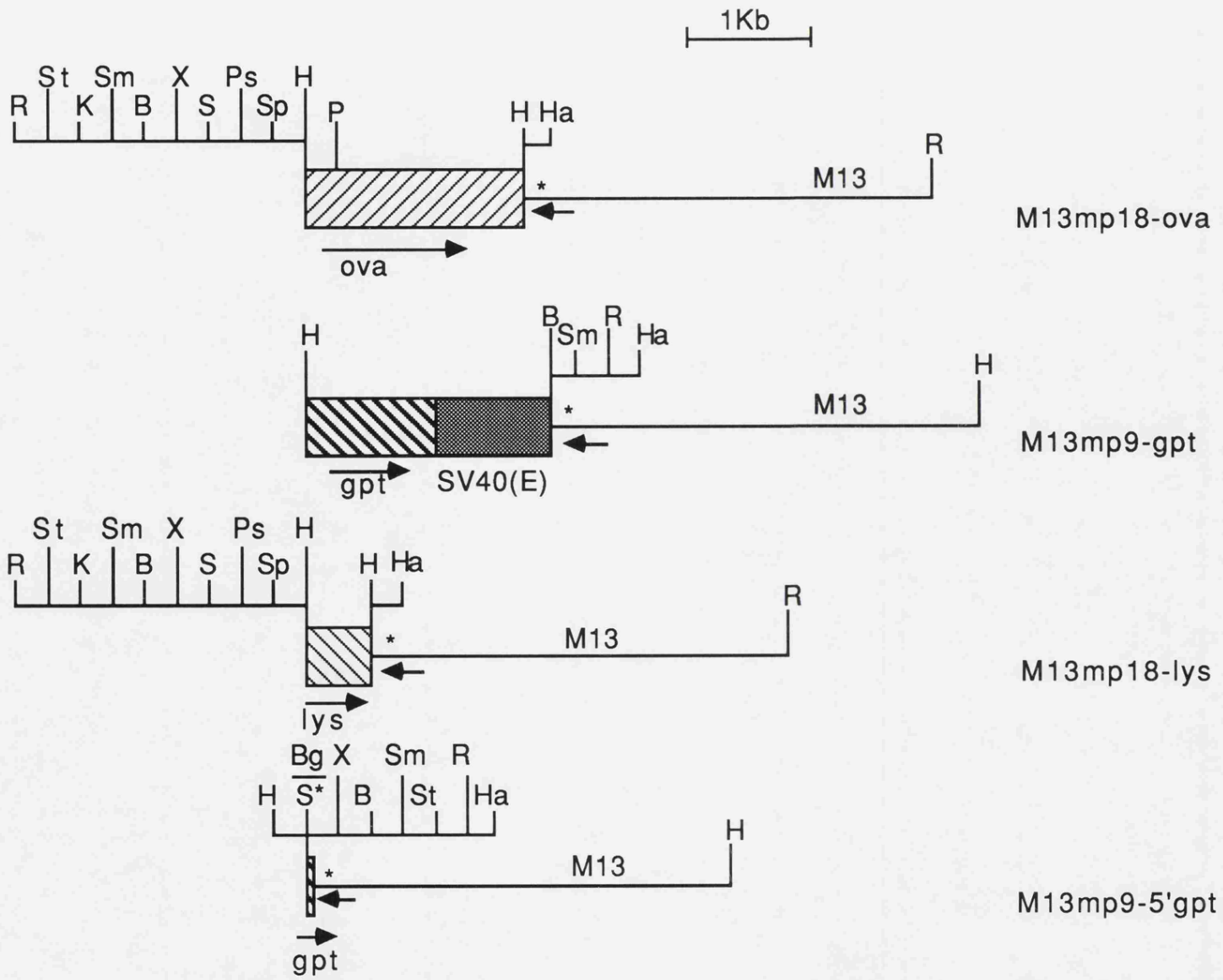
Figure 2.1. M13 recombinants for generation of single-stranded DNA probes and primers.

Partial or complete coding-sequences, for the proteins expressed in this project, were inserted into the polylinker sequence of either M13mp9 or M13mp18, to produce the M13 recombinants illustrated. The construction of these recombinants is described in section 2.10.1.1. These M13 recombinants are illustrated as linearised at the restriction site indicated, at the end of the polylinker sequence.

M13mp18-ova contains a 1.4Kb ovalbumin-cDNA fragment, which is represented by a box filled with thin diagonal lines (extending from bottom left to top right). M13mp9-gpt contains a 1.9Kb fragment, containing the *gpt* coding-sequence, and downstream sequences derived from SV40 (the small t-antigen intron and the early polyadenylation-signal), which are represented by a box filled with thick diagonal lines and a darkly shaded box respectively. M13mp18-lys contains a 0.5Kb lysozyme-cDNA fragment, which is represented by a box filled with thin diagonal lines (extending from top left to bottom right). M13mp9-5'gpt contains a 0.12Kb *Hind* III - *Bgl* II fragment, derived from the 5'-end of the *gpt* sequence in pSV_2gpt, which is represented by a box filled with thick diagonal lines. The thin lines depict the M13 backbone.

The asterisks above the M13 backbone indicate the region where the single-stranded (+ strand) DNA of the M13 recombinant phage is complementary to the universal primer. The arrows indicate the direction of DNA synthesis from the 3'-end of the primer, using this single-stranded M13 recombinant DNA as a template (see section 1.10.1.3.). Restriction sites are abbreviated as follows: *Eco* RI: R, *Sst* I: St, *Kpn* I: K, *Bam* HI: B, *Xba* I: X, *Sal* I: S, *Pst* I: Ps, *Sph* I: Sp, *Hind* III: H, *Pvu* II: P, *Hae* III: Ha, *Bgl* II: Bg. A *Sal* I site, deleted during construction of M13mp9-5'gpt, is indicated by an asterisk.

M13mp18-ova, M13mp9-gpt and M13mp18-lys were used to make hybridisation probes, for detection of DNA sequences in Southern blots and/or RNA sequences in dot blots (see sections 4.2 and 4.7.). M13mp9-5'gpt was used to generate a primer for primer-extension analysis of *gpt* transcripts (see section 3.6.).


1Kb
St Sm X Ps H
R K B S Sp
P
H Ha
R
M13
*
ova
M13mp18-ova
H
B Sm R Ha
H
M13
*
gpt
SV40(E)
M13mp9-gpt
St Sm X Ps H
R K B S Sp
H Ha
R
M13
*
lys
M13mp18-lys
Bg X Sm R
H S* B St Ha
H
M13
*
gpt
M13mp9-5'gpt

17mer oligonucleotides were synthesised by J. Keyte using phosphotriester chemistry on controlled-pore glass (Sproat and Bannwarth, 1983). The synthetic efficiency was good and resulted in an 80-90% yield of the full-length 17mer oligonucleotides. Therefore, purification of the 17mers was not considered necessary. The DNA provided was simply ethanol-precipitated, and resuspended in TE-buffer (pH8.0). Two primers were used in this study:

Universal primer (Anderson *et al.*, 1980):	5'-GTAAAACGACGGCCAGT-3'.
5'-gpt primer (this study):	5'-GCCTCTCCCTGACGCGG-3'.

2.10.1.3. Preparation of probes / primers.

Universally - ^{32}P-labelled single-stranded DNA fragments, corresponding to insert fragments in the M13 recombinants, were made using a method adapted from that of Hu and Messing (1982).

Single-stranded M13-recombinant DNA (see section 2.9.5.) was thawed from frozen, and then heated at 65°C for 5 minutes. A 17mer-oligonucleotide primer (see section 2.10.1.2.) was then annealed to a complementary sequence on the single-stranded recombinant-M13 DNA template, by mixing the following solutions, and incubating for 45 minutes at 65°C:

M13 single-stranded DNA	0.5µg (≈5µl)
10x TM-buffer (100mM Tris.HCl (pH8.5), 100mM $MgCl_2$)	1µl
17mer oligonucleotide primer (2µg/ml)	2µl
H_2O	to 11µl

This mixture was allowed to cool slowly at room-temperature for 10 minutes, and then the following solutions were added:

CGT mix (0.125mM dCTP/dGTP/dTTPs) 2.5mM Tris.HCl (pH8.0), 25µM EDTA)	10µl
M13 TE-buffer (10mM Tris.HCl (pH8.0), 1mM EDTA)	6µl
α-^{32}P-dATP (≈ 10mCi/ml, 300Ci/mmol)	3µl
Klenow (5 units/µl)	0.7µl

The resultant mixture was incubated at 37°C for 15 minutes. This resulted in Klenow extending from the 3'-end of the primer, and

synthesising a radiolabelled-strand complementary to the M13-recombinant template strand. Then 2.5µl of 'cold' ATP (0.5mM) was added, and the mixture was incubated for a further 15 minutes at 37°C. The cold ATP acted as a 'chase', because ATP was no longer limiting, and this allowed the newly-synthesised radiolabelled strand to be extended more rapidly. Then a restriction endonuclease was added which cut out the fragment corresponding to the insert-DNA; this required the restriction site to be located within the polylinker sequence, and on the side of the insert distal to the sequence which hybridised to the oligonucleotide primer (see figure 2.1.). Also, it was, ideally, a unique site in the M13-recombinant. 10 units of restriction endonuclease, 3µl of the appropriate 10x buffer and 1.2µl of 0.1M spermidine were added, and the mixture was incubated at 37°C for 1 hour.

The radiolabelled strand corresponding to the insert fragment was dissociated from the template strand, and then isolated on a gel. Fragments intended simply for hybridisation analysis (see section 2.11.4.), were dissociated from the template strand by addition of 5.2µl of a 1.5M NaOH / 0.1M EDTA solution, and then isolated by electrophoresis through a low-melting-point agarose gel (see section 2.7.7.2.). Following electrophoresis, the radiolabelled strand was located by autoradiography. The gel was wrapped in cling-film and a piece of Kodak X-omat AR film was layed on top of the gel and exposed for 5 minutes. The developed film revealed that the radiolabelled strand was located between the M13-backbone strand (at the top of the gel) and the unincorporated nucleotides (at the bottom of the gel). A gel slice, containing the strand, was cut from the gel using a scalpel, and stored at -20°C. When required for hybridisations, the gel was melted at 68°C, and the molten agarose-solution was added directly to the hybridisation solution.

For isolation of smaller fragments (e.g. primers for primer-extension), polyacrylamide-gel electrophoresis was used (Maxam and Gilbert, 1980). Following restriction digestion of the M13-recombinant extension-reaction product (see above), the DNA was ethanol-precipitated and the pellet was washed twice in 70% (v/v) ethanol. The pellet was vacuum-dried and dissolved in 10µl of denaturation buffer (30% (v/v) DMSO, 1mM EDTA, 0.05% (w/v) xylene cyanole). This solution was heated to 90°C for 3 minutes, and then flash-frozen on a dry-ice / isopropanol bath. On thawing,

the denatured-DNA solution was quickly loaded onto a non-denaturing polyacrylamide gel. For the isolation of DNA fragments smaller than 200 bases long, an 8% (w/v) acrylamide / 0.24% (w/v) bisacrylamide gel, in 0.5x TBE-buffer (45mM Tris.HCl (pH8.3), 45mM boric acid, 1.25mM EDTA), was used. The gel was made by mixing the following solutions:

30% (w/v) acrylamide / 0.9% (w/v) bisacrylamide, (deionised)	13.3ml
1x TBE-buffer	25ml
TEMED	15µl
10% (w/v) ammonium-persulphate solution (fresh)	50µl
H_2O	11.7ml

Standard sequencing-gel plates and apparatus (see section 2.10.2.2.) were used for running the gel. 0.5x TBE-buffer was used in both the upper and lower reservoirs. The gel was pre-electrophorised, at 1,000 volts for 30 minutes, before loading the samples. The samples were electrophorised at the same voltage until the xylene-cyanole dye had run half way down the gel (≈ 1½ hours). The glass gel-plates were dismantled, leaving the gel on the unnotched (unsiliconised) plate. The radiolabelled strand was located by autoradiography, as described for agarose gels (see above). The DNA was purified from the gel by electroelution, as described in section 2.7.7.2.1. The DNA was then recovered by ethanol-precipitation using 2.5 volumes of ethanol.

2.10.2. DNA sequencing.

DNA sequencing was performed for several reasons; for determining the orientation of the lysozyme-cDNA insert in M13mp18-lys (see section 2.10.1.1.); for ensuring that the annealing of an oligonucleotide primer to a M13-recombinant single-stranded DNA template was specific (see section 3.6.2.); and as a source of a ladder of DNA fragments of known size, sequentially increasing in size by 1 base, for use as markers in gels containing the products of primer-extension reactions (see section 3.6.2. and figure 3.9.).

2.10.2.1. The DNA-sequencing reaction.

DNA sequencing was performed using the dideoxy chain-termination sequencing approach, first described by Sanger *et al.* (1977), but using α-^{35}S-dATP as developed by Biggin *et al.* (1983).

The primer was annealed to a recombinant-M13 single-stranded DNA template (see section 2.10.1.3.). Then 3µl of 10mCi/ml (4000 Ci/mMol) α-^{35}S dATP and 1.5 units (≈ 1µl) of Klenow were added to the annealed template / primer mixture. For each clone being analysed, four microfuge tubes were labelled (A,C,G,T) and 2.5µl of the template / primer / label / enzyme mix was added to each. Then 2µl of a NTP mix labelled 'A', 'C', 'G' or 'T' was added to the rim of the corresponding tube. The NTP mixes were made up by mixing the following solutions (all volumes in µl):

	'A'	'C'	'G'	'T'
0.5mM dCTP	20	1	20	20
0.5mM dGTP	20	20	1	20
0.5mM dTTP	20	20	20	1
M13 TE-buffer (10mM Tris.HCl (pH8.0), 0.1mM EDTA)	20	20	20	20
0.02mM ddATP	80			
0.02mM ddCTP		80		
0.05mM ddGTP			80	
0.5mM ddTTP				80

The reactions were initiated simultaneously by pulse-centrifuging the NTP mixes to the bottom of the tube. Then the tubes were incubated at 37°C for 30 minutes. Then 2µl of cold 0.5mM dATP was added to the rim of each tube. The contents were mixed by pulse-centrifugation, and incubated for a further 15 minutes at 37°C. This incubation acted as a chase. At this stage the procedure could be stopped by putting the tubes on ice, and then storing at -20°C. Before loading on the sequencing gel, 3µl of formamide-dye mix (made by mixing 100ml formamide (previously deionised with 5g of Amberlite mixed-bed resin (MB1)), 0.03g xylene cyanole, 0.03g bromophenol blue and 0.75g Na EDTA) was added to each tube, and the tubes were placed in a 90-95°C water bath for at least 3 minutes.

2.10.2.2. Electrophoresis of sequencing-reaction products on polyacrylamide gels.

Resolution of dideoxy-sequencing reaction products was achieved by vertical electrophoresis through 6% (w/v) acrylamide / 48% (w/v) urea, buffer-gradient gels, as developed by Biggin *et al.* (1983).

Two stock acrylamide-solutions, top-gel and bottom-gel mixtures, were prepared by mixing the following solutions:

	Top mix.	Bottom mix.
38% (w/v) acrylamide / 2% (w/v) bis., (deionised)	24ml	4.5ml
Urea	76.8g	14.4g
10x TBE-buffer	8ml	7.5ml
Sucrose	-	3.0g
Bromophenol-blue solution (0.01g/ml)	-	0.3ml
H_2O	to 160ml	to 30ml

These stock solutions were stored in the dark at 4°C. Immediately before pouring the gel, solutions A and B were made by mixing the following solutions:

	Solution-A	Solution-B
Top-gel mixture	60ml	-
Bottom-gel mixture	-	15ml
10% (w/v) ammonium-persulphate soln. (fresh)	360µl	90µl
TEMED	15µl	5µl

8ml of solution-A, followed by 12ml of solution-B were drawn up into a 25ml glass-pipette using a Drummond pipet-aid, and a partial buffer-gradient was formed by introducing three air-bubbles into the pipette to mix the interface. The gradient was pipetted between two taped glass-plates. These plates (34 x 40 cm), one of which was siliconised, had been taped together after separating with 0.4mm spacers. The gradient was added along one side of the plates, with the plates inclined at an angle of 45° to the horizontal. The angle of the plates was then decreased, and the remaining solution-A was added, taking care not to disturb the gradient. The gradient could be seen visually by the increasing density of bromophenol blue towards the bottom of the plates. The plates were laid horizontally, and BRL sharks-tooth combs were inserted (smooth surface into the gel). The plates were clamped at their edges using bulldog-clips, and the gel was allowed to set for at least two hours.

The tape was removed from the edges of the plates and the gel assembly was transferred to a vertical electrophoresis-unit. 0.5x TBE-

buffer was placed in the upper reservoir and 1x TBE-buffer was placed in the lower reservoir. The combs were removed, and then reinserted, this time with the sharks teeth pointing into the gel, so as to form loading-wells. Formamide-dye mix (see section 2.10.2.1.) was added to one of the wells, and then the gel was pre-electrophorised at 1,500 volts (≈ 30mA) for 30 minutes, to ensure the electrophoresis set up was operating correctly.

Immediately before sample loading, the wells were rinsed out using the buffer in the reservoir in order to remove concentrated urea which leached from the gel. 4µl aliquots of the boiled sequence-reaction / formamide-dye solutions (see section 2.10.2.1.) were loaded into adjacent wells, in the order A,C,G,T, (or A,G,C,T) for each recombinant, using a drawn-out glass capillary-tube. The gel was run at 1,500 volts until the bromophenol-blue marker was approximately 6cm from the bottom of the gel.

Following electrophoresis, the gel plates were removed from the electrophoresis assembly, and then prised apart, leaving the gel adhering to the unsiliconised plate. The gel was fixed by immersing in fixing solution (10% (v/v) glacial acetic-acid, 10% (v/v) methanol) for 10 minutes. The fixed gel, still present on the glass plate, was removed from the fixing solution, and then transferred to a sheet of Whatman 3MM paper by pressing the paper firmly down onto the gel. The gel was dried for 2 hours at 80°C using a Bio-Rad model-483 slab-gel dryer, and then autoradiographed for 16-64 hours using Fuji RX or X-omat AR X-ray film.

2.11. Studies on genomic DNA.

2.11.1. Isolation of genomic DNA from myeloma cells.

Genomic DNA was prepared from stably-transfected myeloma cells, in order to identify the presence, and investigate the integration pattern, of exogenous DNA in the host genome.

Typically, 2 x 10^7 myeloma cells were harvested by centrifugation. The supernatant was removed, and the cell pellet was resuspended in 1.83ml of SSPE-buffer (180mM NaCl, 10mM Na_3PO_4 (pH7.4), 1.0mM EDTA). The cells were then lysed, by slowly adding 3.33ml of SDS-EB - solution (2% (w/v) SDS, 0.4M NaCl, 100mM Tris.HCl (pH8.0), 40mM EDTA), with stirring. This suspension was incubated at 50°C for 30 minutes. 5.167g of

ultrapure CsCl was added, and the tube was inverted gently until all the CsCl had dissolved. Then 330µl of EtBr solution (5mg/ml) was added, and the resultant solution was pipetted into a Beckman quick-seal tube, taking care to avoid aggregates of EtBr and protein which formed. The centrifugation and the removal of the DNA band were performed as described for large-scale plasmid-DNA purification (see section 2.7.7.2.). EtBr was extracted by mixing with equal volumes of iso-amyl-alcohol (saturated with H_2O and CsCl), until the pink colour was no longer visible. The DNA was precipitated by adding one volume of H_2O and two volumes of isopropanol. The DNA came out of solution as a string-like precipitate, and this was pelleted by centrifuging, at 12,000 rpm for 1 minute, in a microcentrifuge. The pellet was washed in 70% (v/v) ethanol and then vacuum-dried. The pellet was dissolved in TE-buffer (pH8.0), with heating to 50°C if necessary. The concentration and purity of each genomic DNA preparation was determined spectrophotometrically (see section 2.7.5.).

2.11.2. Digestion and agarose-gel electrophoresis of genomic DNA.

Genomic DNA was prepared on CsCl density-gradients (see section 2.11.1.) and, therefore, contained less contaminants than alkaline-lysis prepared DNA. Consequently, it was efficiently digested by restriction endonucleases. Typically, 15µg of genomic DNA was digested with 15 units of enzyme, overnight, in a reaction volume of 30µl, containing the appropriate restriction-buffer (see section 2.7.6.) supplemented with 4mM spermidine. The extent of digestion was analysed by running a small aliquot on a 1% (w/v) agarose gel (see section 2.7.7.1.); a complete digest produced a smear running from the top to the bottom of the gel.

Once the digest was complete, 5 or 10µg samples of the genomic digests were run on an agarose gel, in preparation for Southern blotting. A 0.8% (w/v) agarose midi-gel was usually used, and size-markers were run alongside. The gel was run overnight (8 volts) to produce good resolution. Then the DNA in the gel was EtBr stained, and photographed with a ruler alongside for subsequent sizing of bands on autoradiograms of Southern blots.

2.11.3. Southern blotting.

Southern blotting involves the transfer of DNA from agarose gels to a membrane solid-support, and was developed by Southern (1975). Following electrophoresis, the agarose gel was trimmed to remove empty tracks. The following protocol (R. Dalgleish, personal communication) was used for a 100cm³ gel; for differently sized gels, the volumes were modified in proportion. The gel was transferred to a tray containing 250ml of 0.25M HCl and agitated gently for 7 minutes. The gel was then transferred, sequentially, into 250ml of denaturing solution (0.5M NaOH, 1.5M NaCl) and then into 250ml of neutralising solution (3M NaCl, 0.5M Tris.HCl (pH7.4)), for periods of 30 minutes each. The gel was rinsed with distilled water between each of these stages.

The blotting apparatus was set up as described by Dalgleish (1987), using 20x SSC-solution (SSC) [3M NaCl, 3M Na citrate (pH7.0)] as the blotting buffer. The gel was transferred to the blotting apparatus, and the edges of the gel were masked with cling-film. A sheet of nylon-filter (Hybond-N, Amersham) and two sheets of Whatman (No.1)-paper were cut to the size of the gel, and pre-wetted with 3x SSC-solution. The nylon-filter was placed on the gel first and overlaid with the Whatman paper. A 5cm pile of paper towels was placed on top of the Whatman paper, and on top of this was placed a glass plate, with a 0.5Kg weight above it. Genomic-DNA blotting was allowed to proceed overnight. However, for the first two hours, paper towels were replaced every ten minutes (as they became wetted) with fresh dry ones, in order to improve transfer.

After the blotting period, the blotting set-up was dismantled. The filter was blotted dry, and the orientation of the filter was marked with a pen for later reference. Then the filter was dried thoroughly by leaving in a 56°C oven for 10 minutes. The dried filter was wrapped in Saran-wrap (Dow Chemical Co.) and then placed, DNA-side down, on a UV-transilluminator. The DNA was covalently cross-linked to the filter by exposing to UV-light for 2 minutes. Filters could then be stored, wrapped in Saran-wrap, at 4°C, until required.

2.11.4. DNA hybridisation.

Hybridisations of radiolabelled probes to DNA on the nylon-filters, were carried out in specially-constructed hybridisation chambers.

These possessed screw-down lids, which allowed the chambers to be fully immersed in a water bath. Filters were incubated at 65°C for at least 1 hour in prehybridisation solution (3x SSC, 5x Denhardt's solution (0.12% (w/v) ficoll type-400, 0.12% (w/v) polyvinylpyrrolidone 40, 0.12% (w/v) BSA), 200μg/ml sonicated and denatured herring-testes DNA (boiled for 10 minutes before adding to the mixture), 6% (w/v) PEG 6000, 0.1% (w/v) SDS). The filters were then transferred to 20ml of hybridisation solution (same as prehybridisation solution, except Denhardt's solution was only 2x), which had been pre-incubated at 65°C. Then an agarose-gel slice, containing a ^{32}P-labelled single-stranded DNA probe (see section 2.10.1.3.), was melted at 68°C, and the molten agarose-solution was added to the hybridisation solution. Hybridisations were carried out overnight, at 65°C.

Non-specifically-bound probe was removed by washing. The filters were rinsed four times, and then incubated twice for periods of ten minutes, in filter-washing-solution I (3x SSC, 0.1% (w/v) SDS), at 65°C. As the probes were perfectly complementary to the genomic sequences being investigated, a higher-stringency wash in filter-washing-solution II (0.5 x SSC, 0.1% (w/v) SDS) was then performed, again at 65°C.

2.11.5. Autoradiography.

Filters were blotted to remove excess washing-solution, but were not allowed to dry completely. They were then wrapped in Saran-wrap and exposed to either Kodak X-omat AR or Fuji RX X-ray film, at -70°C, Using X-ograph intensifying screens.

2.11.6. Removal of hybridised probe from filters.

Stripping of a probe from a nylon-filter was possible providing the filter had been kept moist. This involved incubating the filter in 0.4M NaOH at 45°C for 3 minutes, and then incubating, at 45°C for 30 minutes, in a solution containing 0.1x SSC, 0.1% (w/v) SDS and 0.2M Tris.HCl (pH7.5). The stripped filter was autoradiographed to ensure that most of the probe had been removed, and the stripping process was repeated if necessary. Filters could then be hybridised to a second probe, using the same procedure as for the first hybridisation.

2.12. Studies on RNA.

2.12.1. Preparation of glassware and solutions for RNA work.

To prevent introduction of ribonucleases into solutions used for RNA work, extra precautions were taken in addition to the aseptic technique used when handling DNA (see section 2.7.1.). Whenever possible, RNase-free disposable plasticware was used. Glassware was specifically set aside for RNA work only. Before use, all glassware and non-sterile plasticware was siliconised using a solution of 2% (v/v) dichlorodimethylsilane in 1,1,1-trichloroethane (see Maniatis *et al.*, 1982). Then the plasticware was autoclaved, and the glassware was baked at 300°C, overnight. Solutions were prepared using specially reserved chemicals. They were treated with 0.1% (v/v) diethylpyrocarbonate (DEPC) overnight, and then autoclaved before use.

2.12.2. Isolation of total RNA from myeloma cells.

The ability of LiCl / urea solutions to precipitate RNA was utilised in a protocol obtained from M. Webb (personal communication). Typically, 4 x 10^7 myeloma cells were harvested by centrifugation. The cell pellet was washed, by resuspending in 10ml of DMEM and then recentrifuging. The cells were then resuspended in 6ml of a freshly-made lysis buffer (6M urea, 3M LiCl, 20µg/ml heparin, 10mM Na acetate (pH5.6), 0.1% (w/v) SDS), and transferred to a glass homogenising tube. The cells were homogenised, by manually forcing a teflon pestle repeatedly to the bottom of the tube, until a homogeneous solution of low viscosity was produced. This solution was stored at 4°C, overnight, to allow the RNA to precipitate. The solution was then transferred to a 15ml Corex-tube and centrifuged at 10,000 rpm, at 4°C for 30 minutes, in a MSE Hi-Spin (HS)-21 centrifuge. The pellet was washed, by vortexing thoroughly in 10ml of 6M urea / 3M LiCl solution and then recentrifuging. The pellet was dissolved in 5ml of 0.2M Na acetate (pH5.6) / 0.1% (w/v) SDS solution and then transferred to a glass universal-bottle. This solution was extracted twice with equal volumes of phenol (see section 2.7.3.), with continuous shaking for 15 minutes each time. After each extraction, the phases were separated by centrifuging, at

3,000 rpm for 20 minutes, in a MSE bench-centrifuge. Then a chloroform/iso-amyl-alcohol - extraction was performed, in the same way as for the phenol-extractions. The RNA was precipitated, by adding Na acetate (pH5.6) to a final concentration of 0.2M, then adding 2.5 volumes of ethanol and freezing at -70°C. The RNA precipitate was pelleted in a 30ml Corex-tube, by centrifuging at 10,000 rpm, at 4°C for 30 minutes, in a HS-21 centrifuge. The pellet was washed with 70% (v/v) ethanol. The pellet was vacuum-dried and dissolved in 20µl of water. A typical yield of RNA using this method was 0.3mg of total RNA.

2.12.3. Isolation of poly(A)$^+$ RNA from total RNA.

Poly(A)$^+$ RNA was isolated from total RNA using oligo(dT)-cellulose, as developed by Edmonds *et al.* (1971) and Aviv and Leder (1972). In high-salt buffer, oligo(dT) binds only to the complementary poly(A) tail of poly(A)$^+$ RNA. Therefore, poly(A)$^-$ RNA can be eluted from oligo(dT)-cellulose in high-salt buffer, and then the poly(A)$^+$ RNA can be released in low-salt buffer or water. Isolation had to be performed on several RNA samples, isolated from different transfectant clones (see section 3.6.). Therefore, column chromatography was a laborious approach for elution of poly(A)$^-$ RNA from the cellulose followed by elution of the poly(A)$^+$ RNA. A more efficient approach was that of washing the oligo(dT)-cellulose in large volumes of buffer, and then separating the cellulose from the buffer by centrifuging at room-temperature, at 3,000 rpm for 10 minutes, in a MSE HS-21 centrifuge.

Oligo(dT)-cellulose was regenerated from powder by washing, sequentially, in water, then in 0.1M NaOH / 5mM EDTA solution, and then in water again. It was then equilibriated in loading buffer (20mM Tris.HCl (pH7.6), 0.5M NaCl, 1mM EDTA, 0.1% (w/v) SDS). The RNA sample (typically 300µg in 20µl) was diluted to a volume of 1ml with loading buffer, and then the resultant solution was heated to 65°C for 10 minutes, to denature any double-stranded RNA. The sample was allowed to cool to room-temperature and added to 0.5g (dry-weight equivalent - an excess) of equilibriated oligo(dT)-cellulose. Poly(A)$^-$ RNA was washed from the cellulose (using centrifugation as described above) with four 10ml volumes of loading-buffer, or until the OD_{260nm} of the supernatant was negligible. The

poly(A)$^+$ RNA was then eluted from the cellulose with four x 1ml washes using water, and recovered by ethanol-precipitation.

2.12.4. Primer-extension analysis of poly(A)$^+$ RNA.

Primer-extension analysis was performed in order to determine the 5'-end point(s) (i.e. the transcription-initiation site(s)) of *gpt* mRNA from a series of *gpt* expression-plasmids (see section 3.6.).

A single-stranded universally-radiolabelled primer was made, using single-stranded M13mp9-5'gpt DNA-template (see section 2.10.1.). The amount of radiolabel in the probe was determined by scintillation counting. Typically, 1-5µg of the poly(A)$^+$ RNA was added to an aliquot of the primer (≈ 5,000 cpms), and the mixture was ethanol-precipitated with 2.5 volumes of ethanol. The pellet was redissolved in 8µl of water, and then 2µl of 5x hybridisation-buffer (2M NaCl, 50mM Pipes (pH6.4)) was added. This mixture was drawn up into a 20µl capillary tube, and then the two ends of the tube were sealed by melting. The tube was then put into a 56°C-oven, and the primer was allowed to hybridise to the complementary RNA overnight. The contents of the tube were then added to 90µl of extension-reaction buffer (90µl of this buffer contained 5µl each of 1M Tris.HCl (pH8.2), 0.2M DTT, 0.12M $MgCl_2$, 10mM dATP/dCTP/dGTP/dTTPs and 2.5µl of actinomycin D (1mg/ml) and 10 units of AMV reverse-transcriptase), and the mixture was incubated for 1 hour at 42°C. Nucleic-acid was recovered by ethanol-precipitation and redissolved in 10µl of formamide-dye mix (see section 2.10.2.1.). The sizes of the primer-extension-reaction products were determined on DNA-sequencing gels (see section 2.10.2.2.), running sequencing-reaction products alongside as size markers. Autoradiographic exposures of at least 1 week, at -70°C, using Kodak X-omat AR X-ray film and a X-ograph intensifying-screen, were necessary for detection of the primer-extension products.

2.12.5. Dot-blot analysis of RNA.

RNA dot-blot hybridisation analysis was performed, in order to determine if myeloma-cell transfectants were producing transcripts encoded by transfected heterologous-genes (see section 4.7.3.).

Serial dilutions of total RNA were made in water. One-fifth

volume of 20x SSC / 0.2% (w/v) bromophenol blue was added to the RNA samples, and then 38% (v/v) formaldehyde was added to a final concentration of 7% (v/v). These RNA solutions were heated to 65°C for 20 minutes, and then applied as dots to Hybond-N nylon-filters, which had been presoaked in 6x SSC. Application of the RNA dots was facilitated by clamping the filter in a 'transblot pack', which was attached to a vacuum source. This allowed RNA to be applied to the filter via wells. Therefore, large (30µl) aliquots could be added, and diffusion was limited. Each well was then washed twice with 30µl of 20x SSC. The filter was removed from the pack, and the orientation of the filter was marked with a pen. The RNA was cross-linked to the filter, prehybridised, hybridised to the appropriate probe and then washed, using the procedure described for hybridisation to Southern-blots (see section 2.11.4.).

CHAPTER 3: ASSESSMENT OF ACTIVITY OF EXPRESSION ELEMENTS IN J558L.

3.1. Introduction.

The overriding factor governing the level of protein expression from most protein-coding genes is the rate at which the gene is transcribed. The rate of transcription depends to a large extent on the activity of *cis*-acting sequences called promoters and enhancers, which control the rate of transcription initiation by RNA-polymerase II (see section 1.9.). The aim of this project was to develop a highly-efficient expression system for heterologous genes in the myeloma cell-line, J558L. Consequently, one of the most important tasks was to identify *cis*-acting expression elements with high-level activity in this cell-line. These would then be incorporated adjacent to heterologous coding-sequences in expression plasmids (see chapter 4). The rational was that following transfection of such plasmids into J558L, high-level transcription of the heterologous genes would be obtained and, therefore, give the potential for high-level expression of the heterologous proteins.

The most common approach used for comparing the relative activity of different expression elements in any cell-type, is to use transient assays. Expression plasmids are constructed which contain the same test ('reporter') coding-sequence, but which contain different expression elements in the flanking regions of the coding sequence. Usually, 48-72 hours following transfection, the relative levels of expression from the test sequence in each plasmid are compared. Assays may measure RNA levels, but an increasingly popular approach is to measure the level of protein expression, which usually reflects the level of mRNA. For example, the determination of the level of chloramphenicol-acetyl-transferase (CAT) activity in a mammalian cell following transfection of the bacterial *cat*-gene is commonly used, because the assay is simple and there is no endogenous CAT-activity in mammalian cells (Gorman *et al.*, 1982a). Such transient activity is often a good indication of the average level of expression that can be expected when the same DNA sequences become integrated into the genome of the host cell. Therefore, this approach has

been widely used to compare expression elements (e.g. Edlund *et al.*, 1985).

In spite of the common use of transient assays, it should be noted that the microenvironment of integrated DNA sequences is likely to be substantially different from the extrachromosomal environment of transiently-present DNA sequences. The genomic DNA is packaged by histone proteins (see section 1.9.1.3.), whereas there is little or no packing of transiently-infected extrachromosomal sequences. This difference may result in different rates of transcription of the same gene in chromosomal and extrachromosomal environments. In addition, the transcription rate of any integrated gene is likely to be affected by secondary modifications made to the chromatin within and flanking the gene (see section 1.9.1.3.). Furthermore, transcription of the genomic DNA may be interrupted by events such as DNA-replication, whereas many transiently introduced plasmids are not replicated. The extent to which transcription is affected by chromosomal integration may be influenced by the gene itself and, therefore, the effect may differ between genes. Hence, it is no surprise that the results of transient-expression analysis on a series of expression plasmids do not always correlate with those of stable-expression analysis on the same series of plasmids (e.g. Yoshimura and Chaffin, 1987).

The aim of this study was to produce stably-transfected myeloma cells giving stable high-level heterologous-protein expression from transfected heterologous-genes. The use of such stable transfectants allows long-term production of recombinant protein in 'continuous culture', and is usually preferable to transient systems. Transient systems necessitate the use of short-term, usually small-scale batch cultures, although they may be of use for producing cytotoxic proteins. For this study, it was decided to use a stable-expression system; passive expression-plasmids (see section 1.5.1.2.) and the transfection procedure of spheroplast fusion (see sections 1.4. and 2.3.1.) were used, because such a strategy had previously given rise to stable transfectants in myeloma cells (Oi *et al.*, 1983). Passive expression-plasmids depend on integration into the host genome for their propagation. Hence, and in view of the potential errors associated with the use of transient assays, it was decided to compare the influence of various expression elements on gene expression in stable transfectants, rather than in transiently infected cells.

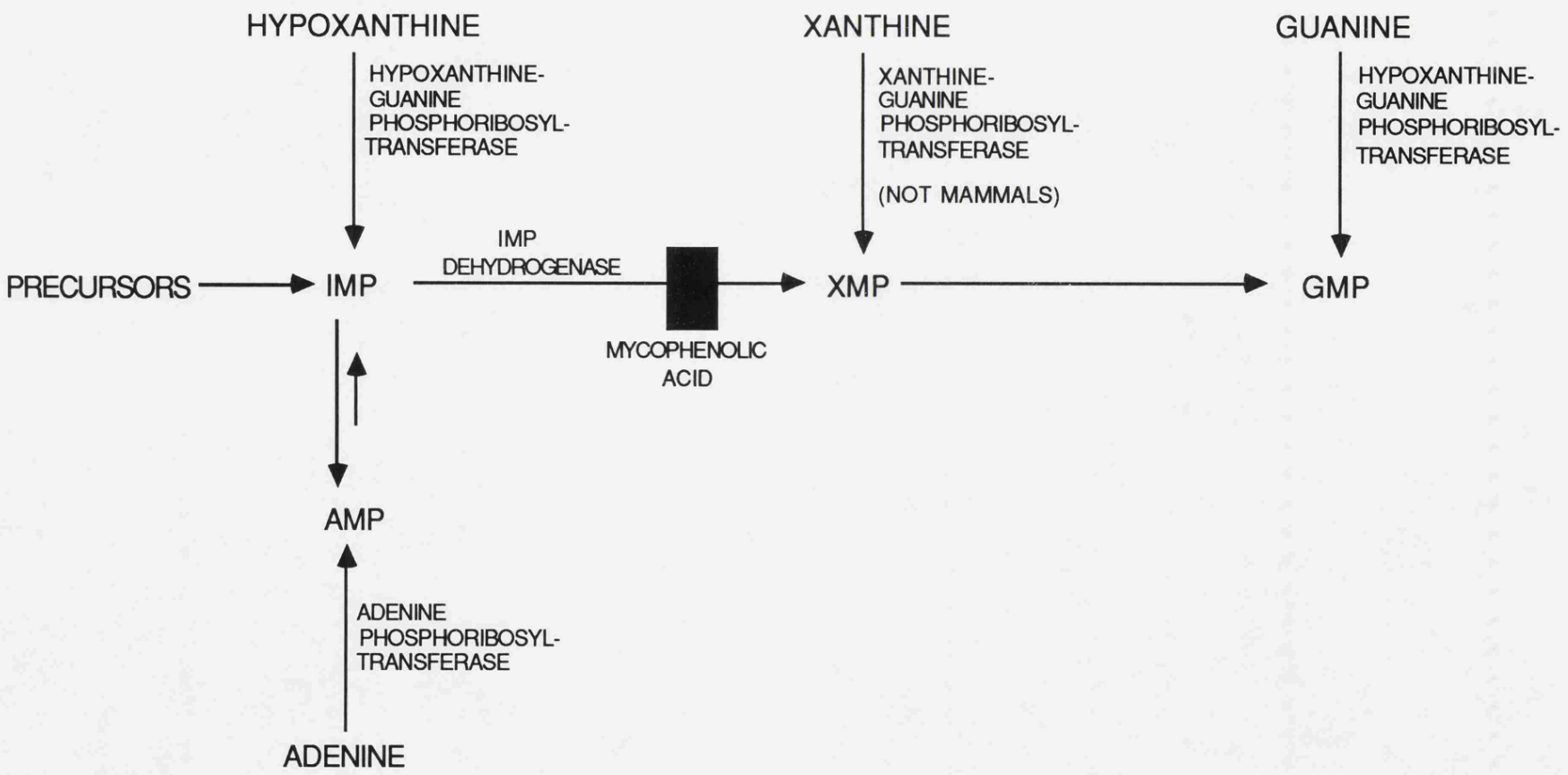
As alluded to above, at the outset of the project the best

developed transfection procedure for introducing exogenous DNA into myeloma cells was spheroplast fusion (Oi *et al.* 1983). Stable-transfection frequencies using this method are rarely greater than 1 x 10^{-3} stable transfectants per cell. The low stable-transfection frequencies are thought to partially reflect the fact that only a low percentage of transfected-cells stably maintain the exogenous DNA (they also reflect the number of cells expressing the product of a selectable exogenous-gene above a certain level, as discussed in greater detail below). Therefore, it is likely to be difficult to assay or even detect expression from stably-transfected cells amongst a background of non-stably-transfected cells. Recently, new transfection-procedures have been developed, such as electroporation, which can give transfection frequencies greater than 1 x 10^{-2} transfectants per cell (Chu *et al.*, 1987; Margolskee *et al.*, 1988). If the efficiency of the electroporation method continues to improve, it may become possible to assay stable-expression amongst a background of non-expressing cells. This was not, however, an option during the early stages of this study.

Due to the non-expressing background problem, a selection system is normally used to identify cells containing, and expressing from, stably-transfected exogenous DNA-sequences. One system that has been successfully used in myeloma cells (Oi *et al.*, 1983), is that based on the bacterial gene, *gpt*, which encodes the enzyme, xanthine-guanine phosphoribosyl-transferase (XGPRT; 5-phospho-α-D-ribose-1-diphosphate: xanthine phosphoribosyltransferase, EC 2.4.2.22.; Mulligan and Berg, 1980 and 1981). Following transfection with a *gpt* expression-plasmid, cells are grown in medium supplemented with xanthine, hypoxanthine and mycophenolic acid (XHMPA). Mycophenolic acid is an inhibitor of IMP-dehydrogenase, which is an essential enzyme in the endogenous biosynthetic-pathways leading to guanine-nucleotide synthesis. In XHMPA-supplemented medium, expression of XGPRT from the transfected *gpt*-gene allows for dominant selection of stable transfectants; only cells synthesising XGPRT are able to use ('salvage') xanthine to allow synthesis of essential guanine-nucleotides by a pathway that bypasses IMP-dehydrogenase. Consequently, transfected cells synthesising sufficient guanine nucleotides survive in the selective medium and give rise to colonies, whereas cells not expressing XGPRT, or sufficient XGPRT, die. The basis of this selection system is illustrated in figure 3.1. The *gpt* gene was used as a selectable marker for the selection

Figure 3.1. Pathways of purine-nucleotide biosynthesis.

This figure shows the biosynthetic pathways leading to synthesis of purine nucleotides, around which the dominant selection-system for mammalian cell-transfectants using the *E. coli* gene, *gpt*, is based (see section 3.1.). In the presence of mycophenolic acid, normal guanine-nucleotide synthesis in mammalian cells, from hypoxanthine via a 'salvage' pathway, or from precursors via a *de novo* pathway, is inhibited. This is because mycophenolic acid is an inhibitor of the enzyme, IMP-dehydrogenase. However, cells expressing the enzyme, xanthine-guanine phosphoribosyl-transferase (XGPRT), from a transfected *gpt*-gene, are able to 'salvage' xanthine, which is added to the medium, to synthesise guanine nucleotides and allow selective survival.
Abbreviations used are: AMP: adenine mono-phosphate; IMP: inositol mono-phosphate; XMP: xanthine mono-phosphate; GMP: guanine mono-phosphate.


HYPOXANTHINE
HYPOXANTHINE-
GUANINE
PHOSPHORIBOSYL-
TRANSFERASE
XANTHINE
XANTHINE-
GUANINE
PHOSPHORIBOSYL-
TRANSFERASE
(NOT MAMMALS)
GUANINE
HYPOXANTHINE-
GUANINE
PHOSPHORIBOSYL-
TRANSFERASE
PRECURSORS
IMP
IMP
DEHYDROGENASE
MYCOPHENOLIC
ACID
XMP
GMP
AMP
ADENINE
PHOSPHORIBOSYL-
TRANSFERASE
ADENINE

frequency. This is because a situation may arise where the expression signals are so active, that most transfected cells express the selectable-marker-gene product above the threshold level. In such a situation, a further increase in the activity of the expression signals is not likely to significantly increase the chances of a cell giving rise to a colony. Of course, if a plasmid can be identified which gives higher transfection-frequencies than any other plasmid, this suggests that the correlation holds at the levels of expression from the plasmids being compared. This is what was observed with the plasmids discussed in this chapter.

Stable-transfection-frequency assays have been used to compare expression from different constructs for many purposes. Examples include the identification of BPV expression-elements (Campo *et al.*, 1983), the study of cell-type specificities of enhancers (Yoshimura *et al.*, 1987) and the determination of the effect of introns on thymidine-kinase expression (Lewis, 1986).

As already alluded to, the experiments reported in this chapter compared the levels of XGPRT expression from a series of *gpt* expression-plasmids, using the stable - gpt^+-transfection - frequency assay described above. All the plasmids compared (except in sections 3.7. and 3.8.) contain different transcriptional elements upstream of the *gpt* gene, but are otherwise identical. It is, therefore, most likely that differences in stable gpt^+-transfection frequencies reflect different rates of transcription initiation. However, because the stable-transfection-frequency assay is only an indirect measurement of expression, it can not be discounted that different expression levels from different plasmids are due to one of the many other factors that can influence expression (see sections 1.8. and 1.9.). Identification of such phenomena requires lengthy investigations. For example, to imply that different XGPRT expression-levels are due to different 5'-untranslated regions of *gpt* RNA affecting its stability and/or transport, requires the demonstration that the ratio of mRNA (determined by Northern analysis) to the rate of transcription (determined by nuclear run-off analysis) is different depending on the expression plasmid. To imply that differences in XGPRT levels are caused by differential translation rates of *gpt* mRNA, requires the demonstration that the ratio of mRNA to the amount of protein expressed is different depending on the expression plasmid. The copy number of integrated plasmids may also

affect expression levels. Therefore, in order to ensure that different plasmids do not have different tendencies to become stably integrated into the genome, Southern-blotting analysis on large numbers of transfectants, or representative polyclonal-transfectant populations, is necessary.

Despite the inability to irrefutably explain the results, the stable-transfection-frequency assay offers the opportunity to rapidly compare a large number of upstream elements and obtain a crude indication of their relative activity. It was expected that the elements shown to be the most efficient in this assay would prove to be efficient in subsequent constructs. Even if the results from the assays did not reflect differential transcription-initiation activity, it was still hoped that the sequences responsible for high-level expression would retain their influence on expression from other heterologous genes in similar expression plasmids.

3.2. Technical aspects of the stable-transfection-frequency assay.

The theoretical aspects of the stable-transfection-frequency assay were discussed in section 3.1. In this section, the validity of the approach from a technical aspect is discussed.

For the stable-*gpt*$^+$-transfection-frequency assays, the DNA was introduced into the cells using the spheroplast-fusion transfection method (see section 2.3.1.). Ideally, the same number of copies of *gpt* expression-plasmids are added in each transfection experiment, so that the transfection-frequency produced reflects properties of the plasmid rather than its abundance. This is much easier to control using the alternative transfection procedure of electroporation (see section 2.3.2.), in which a defined amount of purified DNA is added to the cells. However, electroporation was not a readily available transfection method at the outset of this project.

All the spheroplast-fusion transfections described in this chapter were performed with spheroplasts produced from the same bacterial-cell host, TG1, and the growth of TG1 prior to transfection was carefully reproduced (see section 2.3.1.). Also, all the plasmids have the same bacterial origin-of-replication in the pBR backbones. Therefore, it was considered likely that the expression plasmids would be maintained in the

bacterial host at similar copy numbers. It is unlikely that the copy number is limited by the size of the plasmid, because if the copy number is similar to that of pBR322 (≈ 15-40 copies per cell), the plasmid is unlikely to significantly drain the resources of the host cell. Counter transcription (see section 5.1.) may also affect copy number, but this is also unlikely to be a problem because all the expression plasmids used contain only one procaryotic transcription-unit (the ampicillin-resistance gene). Even so, the rules governing copy-number determination are not well understood and, therefore, any DNA sequence might affect copy number in an unanticipated fashion. For example, a sequence containing part of the promoter and some of the coding region of the tetracycline gene in pBR322 causes an unusual supercoiling-distribution (Pruss and Drlica, 1986), and the extent of supercoiling may affect plasmid replication (Gellert, 1981). Therefore, following the chloramphenicol amplification-step, but before making the spheroplasts, for a series of different transfections (see section 2.3.1.), equal aliquots of the bacterial cultures were taken and used to isolate plasmid DNAs as quantitatively as possible, using the alkaline-lysis method (see section 2.7.2.1.). These plasmids (whose structures are described in chapter 3 or 4) differ, with respect to both size and sequence, to a much greater extent than the relatively minor differences between plasmids compared in this chapter. Relative quantitation of the plasmids on an agarose gel (see figure 3.2.) suggested that the copy number varies only to a small extent. Hence, the range of transfection frequencies observed in this chapter are unlikely to be accounted for by different plasmid-copy-numbers in the bacterial host.

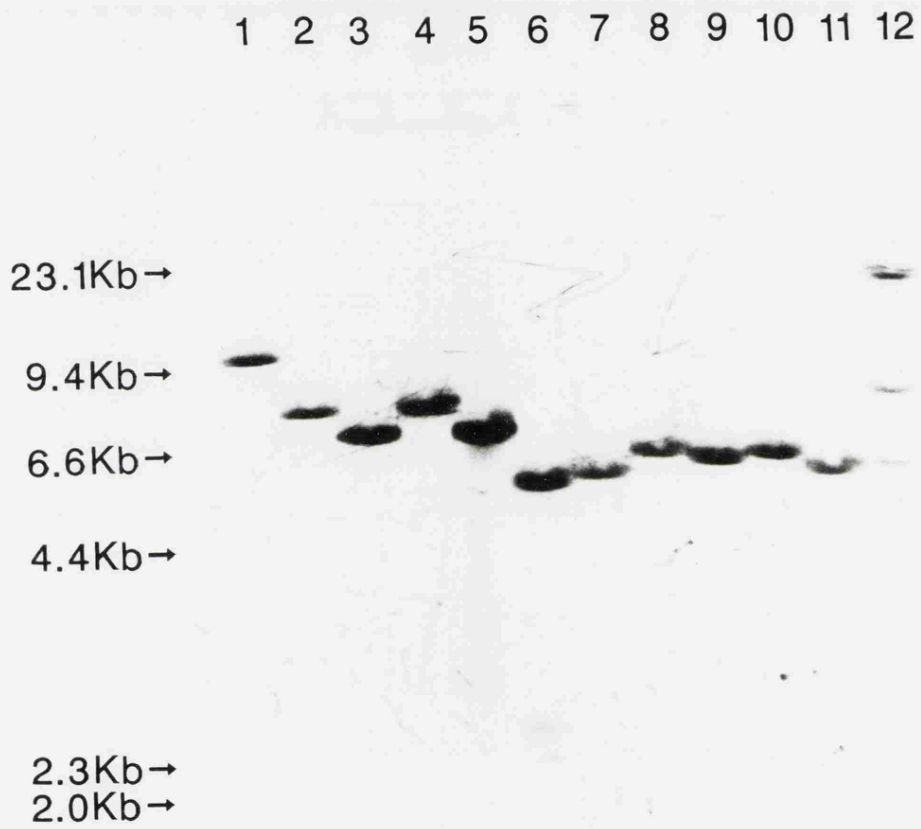
The above result suggests that the copy number of different plasmids in spheroplasts is reasonably constant. If so, then in order for the same amount of DNA to be introduced into J558L in each transfection, the number of spheroplasts used has to be constant. The spheroplasts are toxic to the myeloma cells. Hence, a number of spheroplasts are used such that the viability of myeloma cells following the transfection is between 25 and 50%. Therefore, transfection viability acts as an indication of the reproducibility of spheroplast preparations, because if the number of spheroplasts varies greatly, this will result in a myeloma-cell viability outside the above range. In fact, I found the reproducibility of spheroplast preparations to be good, providing the preparations were made

Figure 3.2. Reproducibility of DNA input into transfection experiments with different expression plasmids.

Throughout this study, the stable *gpt*$^+$-transfection frequencies produced, following transfection with various plasmids, was used as a measure of the influence of flanking sequences in the plasmids on expression from the *gpt* sequence. Transfections were performed by spheroplast fusion, and a similar number of spheroplasts were used in each transfection (see section 3.2.). Hence, for the same amount of *gpt* sequence to be present in each transfection, such that differences in stable *gpt*$^+$-transfection frequencies reflect properties of the plasmids, rather than the number of plasmids introduced in the transfection, requires that the spheroplasts contain similar copy-numbers of each plasmid.

Eleven plasmids were isolated from bacteria which had been grown in preparation for transfections (i.e. chloramphenicol treated, see section 2.3.1.). Their comparative copy-number was then estimated, by running *Eco* RI - linearised fragments on a 1% agarose gel and comparing the intensity of fluorescence from the EtBr-stained fragments (see section 3.2.). The tracks contain the following plasmids: track 1: plysμ; track 2: plysμ2; track 3: plysμA-Δ1; track 4: plysμA-Δ2; track 5: plysμA-Δ3; track 6: pΔTK-Egpt; track 7: plys-SV40; track 8: plys-ISV40; track 9: plys-TK; track 10: plys-βglo; track 11: plys$^+$(d); track 12: *Hind* III - digested λ markers. The plasmids in tracks 1-5 and 7-11 are described in chapter 4, and the plasmid in track 6 is described in chapter 3.

Although there is some variability in band intensities, the largest difference is only approximately five-fold, and generally it is rather less than this. These differences in plasmid copy-number are, therefore, unlikely to account for the large differences (orders of magnitude) in stable *gpt*$^+$-transfection frequencies obtained with some of the plasmids used in this study. Hence, the differences in transfection frequencies must reflect properties of the plasmid sequences rather than simply their relative abundance.


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

on ice (see section 2.3.1.). Therefore, a fixed volume (0.8ml) of spheroplast suspension was always used in the transfections.

Different transfection frequencies are often obtained from separate transfection experiments using the same plasmid; this phenomenon is called transfectional variation. Spheroplast-preparation variability may be partially responsible for this. However, two independent (separate bacterial cultures) transfections, performed on the same day, with the same plasmid (pML-Egpt), gave precisely the same stable *gpt*$^+$-transfection frequency of 1.2 x 10^{-4} transfectants per cell. This result is statistically insignificant, and more experiments are needed to determine how closely transfection frequencies can be reproduced when transfections are performed at the same time. However, I do not feel that experimental variability in the spheroplast-fusion transfection method is a major source of transfectional variation (see below). To categorically demonstrate this, transfections could be performed as cotransfections with an easily assayable plasmid. For example, a plasmid based around pSV$_2$CAT (Gorman *et al.*, 1982), appropriately modified for activity in myeloma cells, could be cotransfected. The CAT activity could then be determined in an aliquot of transient transfectants, while selection for stable *gpt*$^+$-transfectants is made on the remaining cells. If the CAT activities are similar following each transfection, this suggests that all cells are also infected with similar numbers of *gpt* expression-plasmids and, therefore, validates the comparison of transfection frequencies.

Despite the care taken to minimise transfectional variation, stable-transfection frequencies ranging over more than a full order of magnitude were seen following transfections performed with the same plasmid (although, generally, variation was less than this). However, these transfections were performed at different stages of the investigation. Consequently, transfectants were selected for in different batches of medium. It is well known that cloning efficiencies show large variation in media containing different batches of serum. Therefore, even though a transfected cell may express sufficient XGPRT in order to survive selection, it still only has a certain percentage chance of giving rise to a colony, which is dependent on the nutritious value of the medium. I believe this was the major source of transfectional variation, rather than variation in the conditions of the actual transfection-experiment. In view

of this problem, the transfection frequencies listed in tables 3.1., 3.2., 3.3. and 3.5. (but not table 3.4.) were obtained following transfection and selection in media which gave 'normal' (≈ 10^{-4} transfectants per cell) stable *gpt*$^{+}$-transfection frequencies with pΔTK-Egpt (see section 3.3.1.). In later experiments (see chapter 5), when low stable *gpt*$^{+}$-transfection frequencies were obtained with some plasmids, a 'positive control' transfection with pΔTK-Egpt was performed on the same day, to ensure that the low transfection-frequency was due to the properties of the plasmid rather than the cloning potential of the selective medium. Also, the transfection-frequency determinations for many of the plasmids compared in this chapter were repeated, to obtain an average transfection-frequency, so as to reduce errors caused by transfectional variation.

It must be emphasised that the aim of the experiments described in this chapter was to obtain a crude indication of the relative activity of different upstream expression-elements, and combinations of elements, in J558L, rather than identifying small differences in their activity. The transfection frequencies obtained with various classes of constructs are clearly reproducible as discussed in the following sections.

3.3. pPROM-Egpt constructs.

In section 1.12., I described the important role that the IgH-gene enhancer plays in ensuring high-level transcription of the IgH-gene in myeloma cells. It was expected that the presence of this element in expression plasmids would also allow high-level transcription of heterologous genes in myeloma cells and, therefore, give the potential for high-level expression of the encoded product. Hence, a series of pPROM-Egpt constructs (see figure 3.6.) were made, which all contain the mouse IgH-gene enhancer upstream of different promoter, or promoter-like, fragments. These expression signals are positioned upstream of a *gpt* coding-sequence, which is flanked downstream by SV40 early-region sequences (the small t-antigen intron and the polyadenylation signal). All the plasmids contain backbones derived from pBR322, which contain the ampicillin resistance gene. They also contain a fragment derived from the late-region of SV40. Hence, the plasmids are, in effect, derivatives of $pSV_{2}gpt$ (see figure 5.1.), in which the SV40 early-promoter fragment is substituted by

different expression-elements. The construction of these plasmids is outlined below, and the source of all plasmids used in their construction is given in section 2.1.

3.3.1. Plasmid construction.

The first pPROM-Egpt construct to be made was pΔTK-Egpt, and the strategy used for its construction is illustrated in figure 3.3. The starting point was the construct, pTK_2ov, which was conveniently available, and which contains the same *Bam* HI - *Pvu* II backbone fragment as in pSV_2gpt. Inserted into this backbone, is an ovalbumin transcription-unit, consisting of a HSV-TK promoter, an ovalbumin cDNA and a SV40, early polyadenylation-signal. Firstly, an ovalbumin cDNA-fragment was deleted from pTK_2ov by cutting with *Hind* III and ligating the backbone, to produce pTK. Then the ends of a 2.3Kb *Hind* III IgH-gene major-intron fragment, containing the enhancer sequence, were filled-in with Klenow and inserted into the *Pvu* II site of pTK. This IgH-gene fragment was obtained from pXTKVδ6, which is an easier plasmid to manipulate than larger plasmids such as pSV-Vμ1 (see figure 4.4) which contain the same Ig fragment. Only one orientation of the insert was found and this was named pTK-E(a). A 1Kb *Xba* I fragment within this 2.3Kb insert, which has been shown to contain most of the active elements of the enhancer (Banerji *et al.*; Gillies *et al.*; Neuberger, all 1983), was reversed in orientation, by cutting with *Xba* I and then religating the fragments, to produce pTK-E(b'). This manipulation was necessary in order to subsequently place the enhancer elements directly upstream of the promoter fragment (see below). A 1.9Kb *Hind* III - *Bam* HI fragment from pSV_2gpt, which contains the *gpt* coding-sequence, and downstream SV40-sequences including the small t-antigen intron and the early polyadenylation-signal, was then inserted into the *Hind* III and *Bam* HI sites of pTK-E(b'). This manipulation replaced a 21bp translation-termination sequence and the SV40, early polyadenylation-signal, and generated pTK-E(b')gpt. The *Eco* RI site at the junction between the SV40 late-region fragment and the pBR backbone was deleted, by cutting with *Bam* HI and *Acc* I, and then inserting a fragment from a pSV_2gpt derivative (pSV_2gpt-ΔRi) which is identical, except the *Eco* RI site in the fragment is deleted. This allowed for deletion of a remaining 1.4Kb *Eco* RI fragment,

which was carried out by digesting with *Eco* RI and then ligating the backbone. This resulted in removal of IgH-gene sequences which are not essential for enhancer activity, and also removal of the upstream region of the HSV-TK promoter, including some regulatory sequences (McKnight *et al.*, 1981, 1984). Hence, the enhancer was now directly upstream of a truncated TK-promoter (ΔTK). This plasmid was called pΔTK-Egpt.

pΔTK-Egpt is a very convenient plasmid to manipulate because the ΔTK promoter is flanked by unique *Eco* RI and *Hind* III sites. It was, therefore, used as the parent plasmid for constructing all other pPROM-Egpt constructs (see figure 3.6.); the ΔTK promoter was removed by *Eco* RI and *Hind* III digestion, and replaced with a series of other expression-elements which were either already available on *Eco* RI - *Hind* III cassettes, or were inserted into the polylinker sequence of pUC8 to produce such a cassette. These manipulations are outlined below and in figures 3.4. and 3.5.

A 2.1Kb *Hind* III (filled in) - *Bam* HI fragment containing the promoter sequence and the coding sequences for the leader, variable, diversity and J_2 regions (L-VDJ_2) of an IgH gene (the same as in pSV-Vμ1, see figure 4.4) was obtained from the plasmid pXTKVδ1. It was then cloned into the *Sma* I and *Bam* HI sites of the pUC8 polylinker, to generate pUC8-ULVDJ_2. The Ig coding sequences were removed, by cutting with *Nco* I and *Bam* HI, filling-in with Klenow and then ligating the backbone. Unfortunately, this regenerated the *Nco* I site, which contains the native IgH-gene translation-initiation codon. Therefore, the plasmid was reopened with *Nco* I, treated with Bal31 exonuclease as described in section 2.7.9., blunt-ended with Klenow and then religated. The resultant plasmid, pUC8-Ig (see figure 3.4.), was a source of the *Eco* RI - *Hind* III promoter cassette which was used to construct <u>pIg-Egpt</u>.

A mouse mammary tumour virus LTR (MMTV-LTR) - containing fragment (1.4Kb) was removed from pSVMdhfr by *Hind* III digestion. The ends of the fragment were filled-in with Klenow, and then it was cloned into the *Sma* I site of the pUC8 polylinker-sequence, to produce pUC8-MMTV. This was a source of the 1.4Kb *Eco* RI - *Hind* III cassette (see figure 3.5.) which was used to construct <u>pMMTV-Egpt</u>.

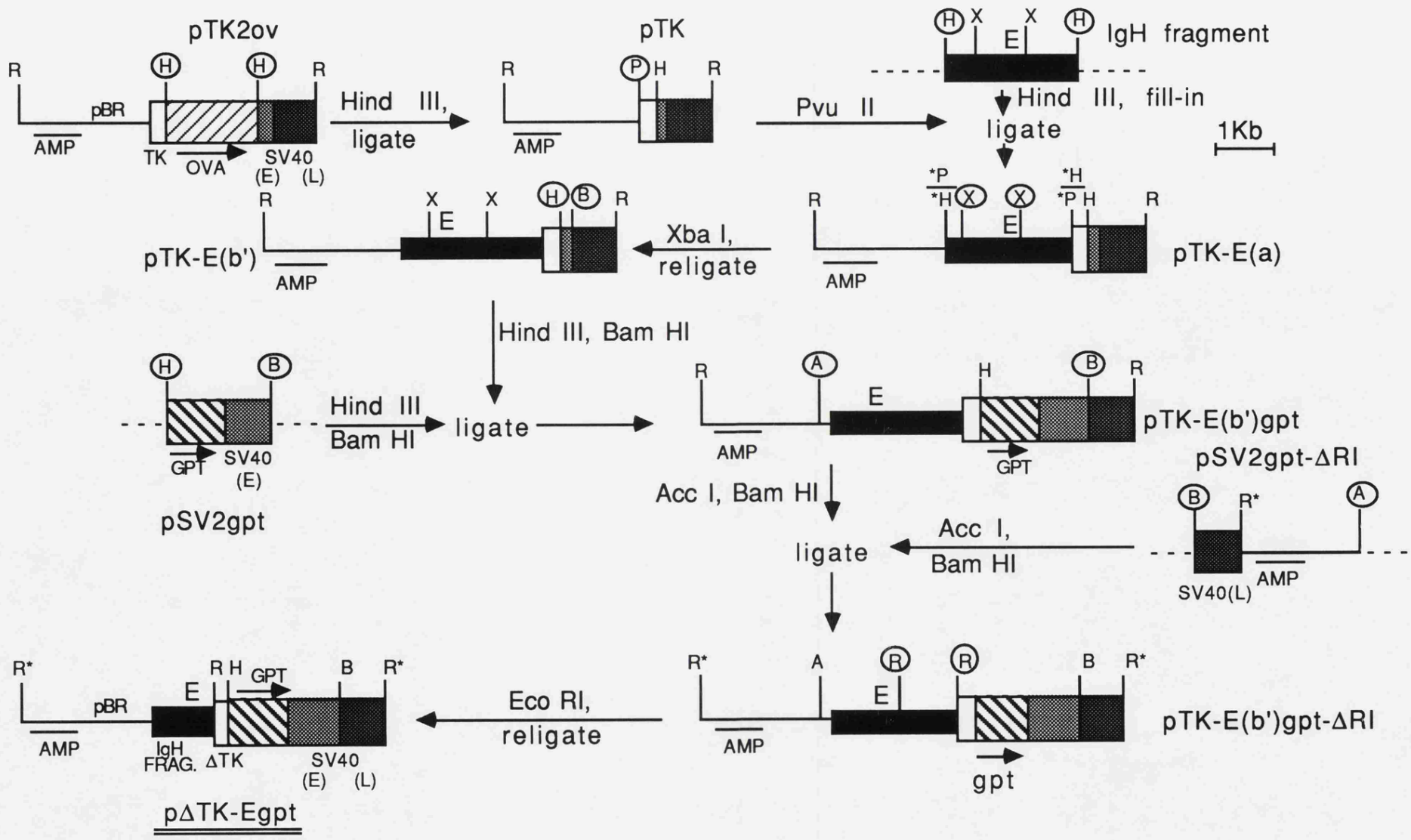
A fragment containing the SV40 early-promoter, which also contains the SV40 enhancer and the SV40 late-promoter (reviewed by McKnight and Tjian, 1986), was removed as a 0.3Kb *Pvu* II - *Hind* III fragment from

Figure 3.3. Construction of pΔTK-Egpt.

The strategy used to construct pΔTK-Egpt is illustrated, and this is described in greater detail in section 3.3.1. Briefly, the construction involved the following manipulations. Firstly, an IgH gene major-intron fragment, containing the enhancer, on a 2.3Kb *Hind* III - fragment, was inserted upstream of a HSV-TK promoter. Then a 1Kb *Xba* I fragment within the IgH sequence, which contains all the enhancer elements, was reorientated. The resulting IgH-enhancer - TK-promoter sequence was used to make a complete *gpt* transcription-unit, by fusion with a fragment from pSV_2gpt, which contains a *gpt* sequence and SV40 sequences downstream. Superfluous Ig sequences and also the 5'-end of the TK promoter were then deleted, to produce pΔTK-Egpt. This plasmid was used as the parent plasmid of all other pPROM-Egpt constructs, because of the unique *Eco* RI and *Hind* III sites flanking the ΔTK promoter (see section 3.3.1.). In addition, pΔTK-Egpt was used in control transfections to indicate the efficiency of transfections and/or the cloning efficiency of selective media (see tables 5.1. and 5.2.).

HSV-TK promoter sequences (the complete sequence (TK), and the deletion derivative (ΔTK)) are represented by unfilled boxes, the *gpt* sequence is represented by a box filled with thick diagonal lines, IgH-gene major-intron fragments, containing the enhancer (E), are represented by thin black boxes and SV40 sequences are represented by darkly shaded boxes (fragments containing the early polyadenylation-signal, with (SV40(E)) or without (SV40(E)pA) the small t-antigen intron, are shaded less darkly than the sequence derived from the late region (SV40(L))). A thin line represents the pBR backbone, which contains the ampicillin-resistance gene (amp).

Plasmids, illustrated in full, are shown linearised at an *Eco* RI site, and for consistency, they are also shown linearised at the same site following deletion of this site. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Pvu* II: P, *Hind* III: H, *Bam* HI: B, *Eco* RI: R, *Xba* I: X, *Acc* I: A. Sites used in each manipulation are circled, for emphasis. Only sites relevant to the manipulations performed are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.


pTK2ov
R
H
H
R
pBR
AMP
TK
OVA
SV40
(E) (L)
Hind III,
ligate
pTK
R
P
H
R
AMP
Pvu II
H X X H
E
IgH fragment
Hind III, fill-in
ligate
1Kb
*P
*H
X
X
*H
*P H
R
R
E
pTK-E(a)
AMP
Xba I,
religate
R
X
X
H B R
E
pTK-E(b')
AMP
Hind III, Bam HI
H
B
Hind III
Bam HI
GPT
SV40
(E)
pSV2gpt
ligate
R
A
H
B
R
E
pTK-E(b')gpt
AMP
GPT
pSV2gpt-ΔRI
Acc I, Bam HI
B
R*
A
Acc I,
Bam HI
ligate
SV40(L)
AMP
R*
A
R
R
B
R*
E
pTK-E(b')gpt-ΔRI
AMP
gpt
Eco RI,
religate
R*
R H
GPT
B
R*
E
pBR
AMP
IgH
FRAG.
ΔTK
SV40
(E) (L)
pΔTK-Egpt

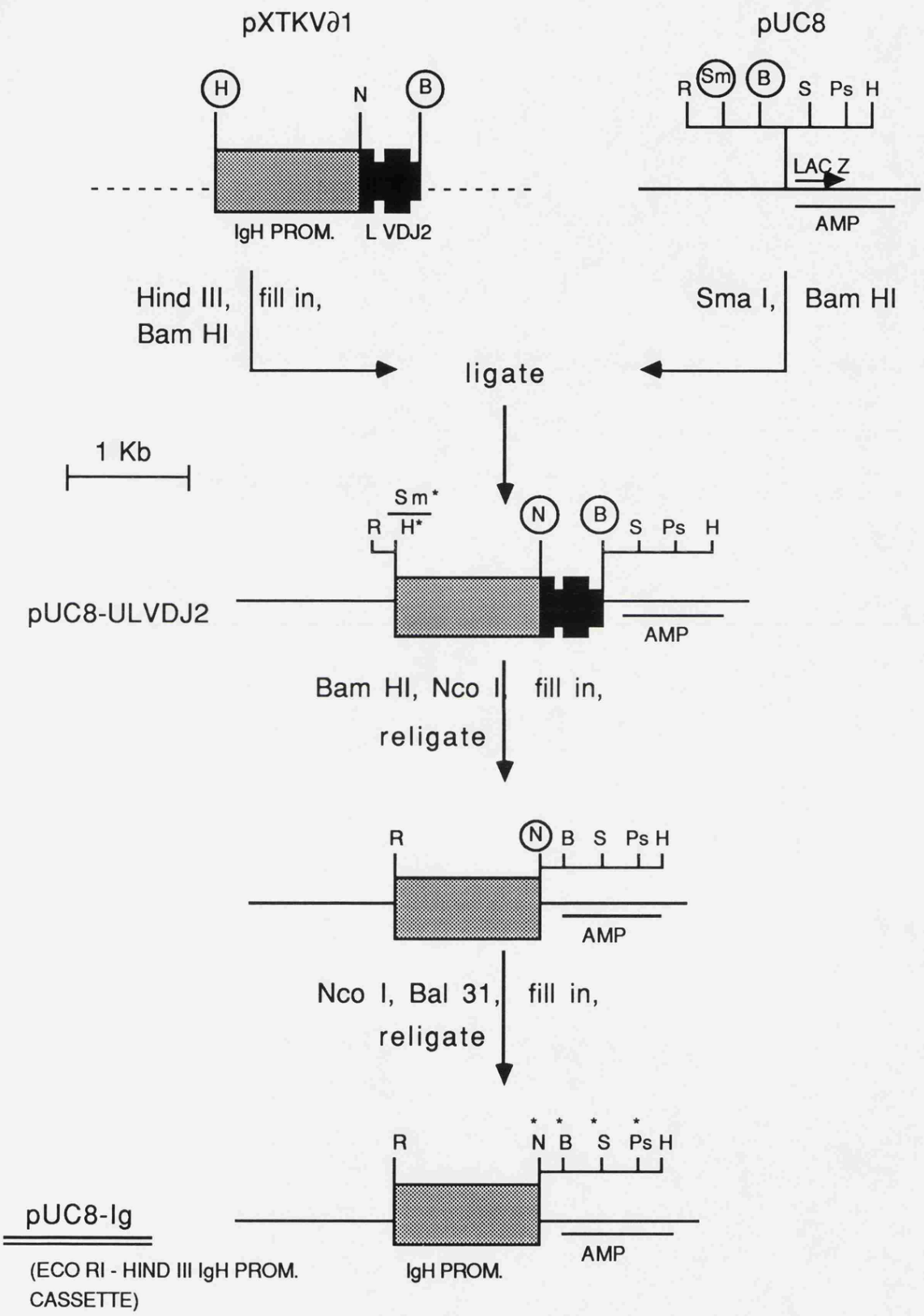

pXTKV∂1
H
N
B
IgH PROM.
L VDJ2
pUC8
R
Sm
B
S
Ps
H
LAC Z
AMP
Hind III,
Bam HI
fill in,
Sma I,
Bam HI
ligate
1 Kb
Sm*
R
H*
N
B
S
Ps
H
pUC8-ULVDJ2
AMP
Bam HI, Nco I,
fill in,
religate
R
N
B
S
Ps
H
AMP
Nco I, Bal 31,
fill in,
religate
R
N
B
S
Ps
H
pUC8-Ig
(ECO RI - HIND III IgH PROM.
CASSETTE)
IgH PROM.
AMP

Figure 3.4. Construction of pUC8-Ig.

The figure illustrates the strategy used to construct pUC8-Ig. This plasmid consists of a 1.4Kb fragment, containing the promoter and upstream region of an IgH gene, inserted into the polylinker sequence of pUC8. The promoter insert is flanked by *Eco* RI and *Hind* III sites, which allowed it to be inserted, in place of the ΔTK promoter, into pΔTK-Egpt, in construction of pIg-Egpt (see section 3.3.1).

Briefly, the construction strategy for pUC8-Ig involved the following manipulations. Firstly, a 2.1Kb IgH-gene fragment (from pXTKVδ1), containing the IgH promoter and variable(V_H)-region coding sequences, was inserted into the pUC8 polylinker. Then the Ig-chain coding sequences were removed in two stages. The coding sequences were deleted almost entirely in the first step, but a *Nco* I site (which contains the native, IgH gene translation-initiation codon (AUG)) was regenerated. Therefore, in order to remove this site, the plasmid was reopened with *Nco* I, and then treated with Bal31 to digest both strands. Then the backbone was blunt-ended with Klenow and religated.

The dot filled boxes represent the IgH promoter and upstream sequences (IgH PROM.). The boxes shaded black represent IgH-gene coding sequences; the wider boxes represent the exons (leader (L) and the actual (contained in the Ig chain) variable-region (VDJ_2)), and the thinner boxes represent the introns. The thin line depicts the pUC8 backbone, and the position of the ampicillin-resistance gene (amp) is indicated.

Plasmids are shown linearised within the pUC backbone, except for pXTKVδ1, for which most of the sequences are not illustrated, and are, instead, represented by a broken line. Restriction sites are abbreviated as follows; *Hind* III: H, *Nco* I: N, Bam HI: B, *Eco* RI: R, *Sma* I: Sm, *Sal* I: S, *Pst* I: Ps. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.

Figure 3.5. Construction of Eco RI - Hind III promoter cassettes containing the MMTV-LTR or the SV40 early-promoter.

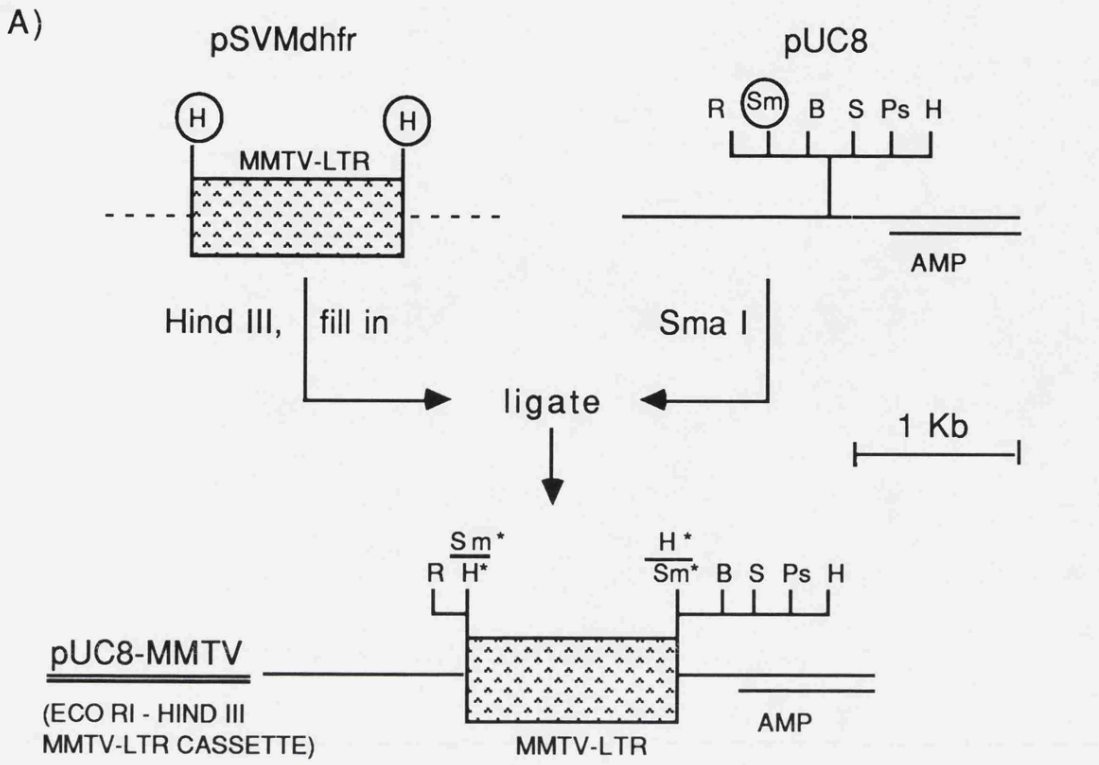
Figure 3.4. illustrated how the IgH promoter was put on an *Eco* RI - *Hind* III cassette, so that it could easily be inserted into the compatible sites of pΔTK-Egpt, the parent pPROM-Egpt construct. Two other promoters were also not immediately available as *Eco* RI - *Hind* III cassettes. The manipulations performed to produce such cassettes are illustrated in this figure.

A) The construction of pUC8-MMTV (see section 3.3.1.) involved cloning of a 1.4Kb MMTV-LTR fragment, with filled-in *Hind* III ends, into the *Sma* I site of the pUC8 polylinker.

B) The construction of pUC8-SV40 (see section 3.3.1.) involved cloning of a 0.3Kb *Pvu* II - *Hind* III fragment from $pSV_{2}gpt$, which contains the SV40 early-promoter (SV40(E)) (and also the SV40 enhancer and late promoter), into the *Sma* I and *Hind* III sites of the pUC8 polylinker.

The MMTV-LTR sequence is represented by a box filled with triangles, and the SV40 early-promoter fragment is represented by a box filled with squares. A thin line represents the pUC backbone, and the position of the ampicillin-resistance gene (amp) is indicated.

The pUC plasmids are illustrated as linearised within the pUC backbone. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows; *Hind* III: H, *Eco* RI: R, *Sal* I: S, *Pst* I: Ps, *Pvu* II: P. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.


A)
pSVMdhfr
H
H
MMTV-LTR
pUC8
R
Sm
B
S
Ps
H
AMP
Hind III,
fill in
Sma I
ligate
1 Kb
Sm*
H*
R
H*
Sm*
B
S
Ps
H
pUC8-MMTV
(ECO RI - HIND III
MMTV-LTR CASSETTE)
MMTV-LTR
AMP

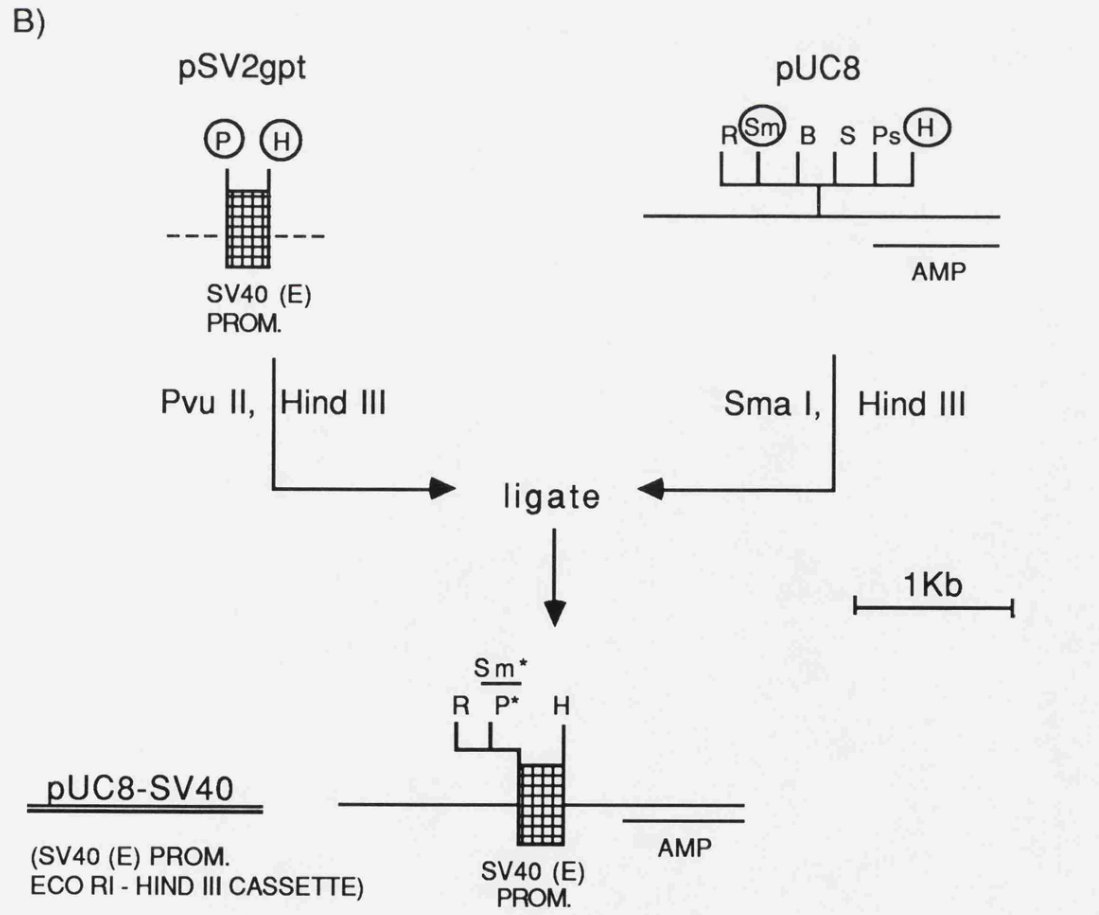

B)
pSV2gpt
P
H
SV40 (E)
PROM.
pUC8
R
Sm
B
S
Ps
H
AMP
Pvu II,
Hind III
Sma I,
Hind III
ligate
1Kb
Sm*
R
P*
H
pUC8-SV40
(SV40 (E) PROM.
ECO RI - HIND III CASSETTE)
SV40 (E)
PROM.
AMP

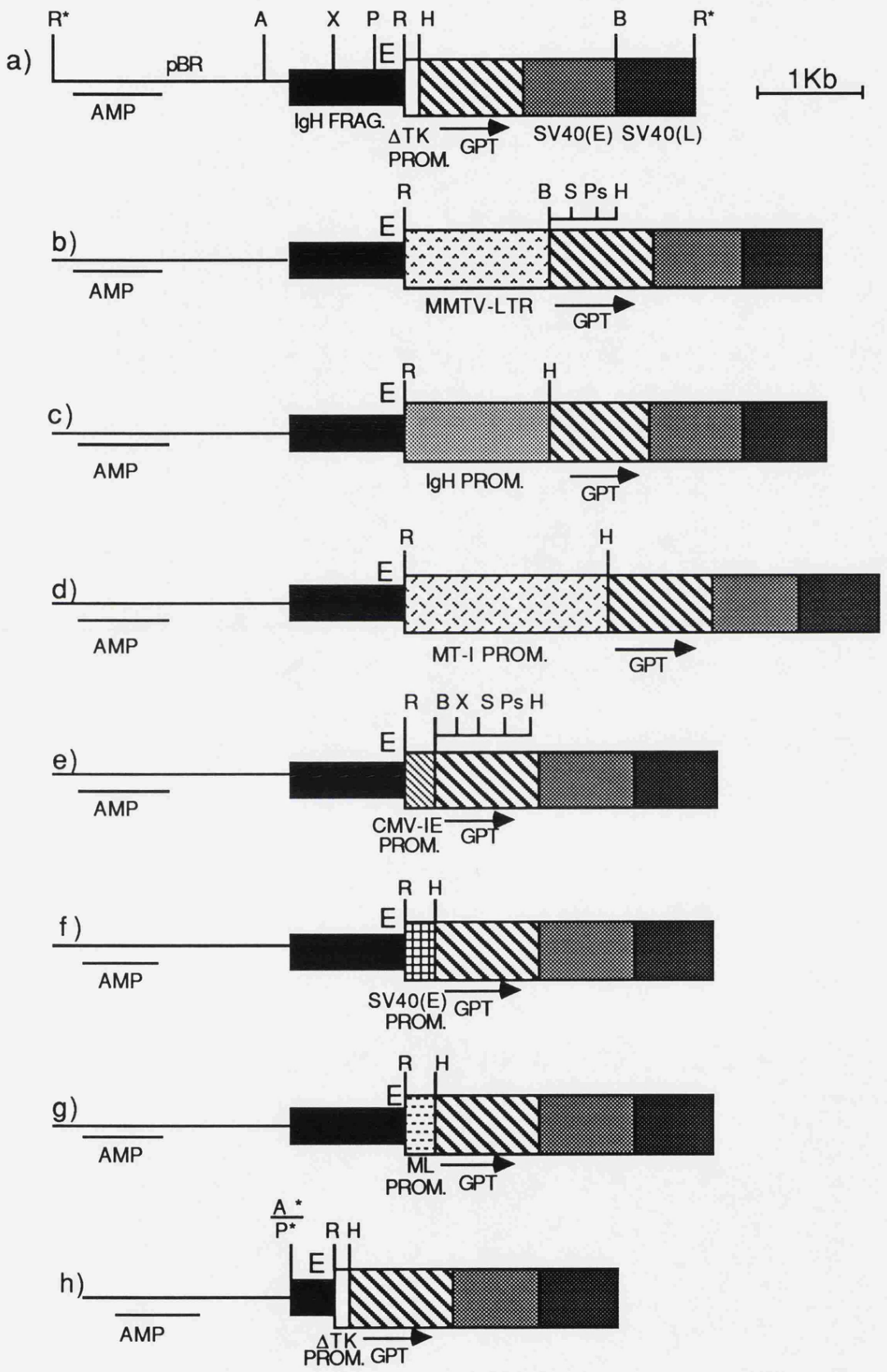
Figure 3.6. pPROM-Egpt constructs.

The structure of all pPROM-Egpt constructs (see section 3.3.1.) is illustrated. Except for pΔTK-E'gpt, they all contain identical *gpt*, SV40, pBR and IgH sequences, and differ only in the promoter fragment directly upstream of the *gpt* coding-sequence. The type of promoter is indicated by the name of the plasmids. The plasmids illustrated are:

a) pΔTK-Egpt.
b) pMMTV-Egpt.
c) pIg-Egpt.
d) pMT-Egpt.
e) pCMV-Egpt.
f) pSV-Egpt.
g) pML-Egpt.
h) pΔTK-E'gpt.

IgH gene major-intron fragments, containing the enhancer (E), are represented by thin black boxes, the *gpt* sequence is represented by a box filled with thick diagonal lines, SV40 fragments downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and early polyadenylation-signal (SV40(E)) is shaded less darkly than the fragment containing sequences from the late region (SV40(L))). The pBR backbone is depicted as a thin line, and the position of the ampicillin-resistance gene (amp) is indicated. Promoter fragments are represented as boxes filled with the following patterns: ΔTK (unfilled), MMTV-LTR (triangles), IgH (dots), MT-I (broken zig-zags), CMV-IE (thin diagonal lines), SV40 early (squares) and Adenovirus major-late (ML) (broken horizontal lines).

The plasmids are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Restriction sites are abbreviated as follows: *Xba* I: X, *Pvu* II: P, *Hind* III: H, *Eco* RI: R, *Bam* HI: B, *Sal* I: S, *Pst* I: Ps, *Acc* I: A. Only a partial restriction map is shown, with emphasis on the sites flanking the promoter fragments. Sites deleted during construction of the plasmids are indicated with asterisks.


a)
R*
A
X
P
R
H
B
R*
E
pBR
AMP
IgH FRAG.
ΔTK
PROM.
GPT
SV40(E)
SV40(L)
1Kb
b)
R
B
S
Ps
H
E
AMP
MMTV-LTR
GPT
c)
R
H
E
AMP
IgH PROM.
GPT
d)
R
H
E
AMP
MT-I PROM.
GPT
e)
R
B
X
S
Ps
H
E
AMP
CMV-IE
PROM.
GPT
f)
R
H
E
AMP
SV40(E)
PROM.
GPT
g)
R
H
E
AMP
ML
PROM.
GPT
h)
A*
P*
R
H
E
AMP
ΔTK
PROM.
GPT

pSV_2gpt, and was cloned into the *Sma* I and *Hind* III sites of the pUC8 polylinker. The resultant plasmid, pUC8-SV40, was the source of the *Eco* RI - *Hind* III promoter-cassette (see figure 3.5.) which was used to construct pSV-Egpt.

The Human Cytomegalovirus immediate-early-region (CMV-IE) promoter was directly available as a 0.3Kb *Eco* RI - *Hind* III fragment from pUC7-CMV cloning plasmid, and this was used to construct pCMV-Egpt.

The mouse metallothionein-I (MT-I) - gene promoter was directly available as a 1.9Kb *Eco* RI - *Hind* III fragment from pMTEB1A, and this was used to construct pMT-Egpt.

The Adenovirus major-late-region (ML) promoter was directly available as a 0.3Kb *Eco* RI - *Hind* III fragment (containing a sequence extending from the *Xho* I site at position 5778 to a *Pvu* II site at position 6071 in the Adenovirus genome) from pMLPCAT, and this was used to construct pML-Egpt.

3.3.2. Results and discussion.

The stable gpt^+-transfection frequencies produced, following transfection of J558L with the various pPROM-Egpt constructs, are presented in table 3.1. The most active plasmid tested was pMMTV-Egpt; it reproducibly gave the highest stable gpt^+-transfection frequency. This plasmid contains the IgH enhancer and the MMTV-LTR, upstream of the *gpt* coding-sequence. However, because of transfectional variation, it is impossible to categorically conclude that this is a superior combination of elements. Indeed, all pPROM-Egpt constructs yielded almost equally-high transfection frequencies. It was initially surprising that pIg-Egpt (which contains an IgH promoter) did not give the highest transfection-frequency, because it has been demonstrated (in transient assays) that the synergism between Ig enhancers and Ig promoters is greater than between Ig enhancers and heterologous promoters (Victor-Garcia, 1986). My results give no indication that the IgH enhancer preferentially activates the IgH promoter. However, this may be due to the size (1.4Kb) of the IgH-gene-derived fragment containing the IgH promoter, which is present in pIg-Egpt. It contains extensive sequences upstream of the promoter, and Doylen *et al.* (1986) showed that extending the size of the IgH-promoter fragment in the

Tables 3.1. - 3.5. Stable-transfection frequencies of J558L to gpt⁺ with different gpt expression-plasmids.

Tables 3.1. - 3.5. list stable *gpt*⁺-transfection frequencies, obtained following transfection of J558L with different *gpt* expression-plasmids. Transfections were performed by spheroplast fusion, and *gpt*⁺ transfectants were selected for in XHMPA selective-medium. The number of colonies produced after 21 days was recorded. The stable-transfection frequency is expressed as the fraction of cells, plated-out following the transfection, that gave rise to colonies.

In some cases, more than one transfection was performed with the same plasmid, and so the range of transfection frequencies is given, and an average transfection frequency is calculated. Most transfections were performed with a single *gpt*-expression-plasmid (S), but some *gpt* expression-plasmids were also used in cotransfections (C) with pTK-E(b')ova (see section 4.2.). In these cotransfections, pTK-E(b')ova was in large excess (approximately eleven-fold more spheroplasts carrying this plasmid than spheroplasts carrying the *gpt* expression-plasmid). Therefore, the transfection frequency was multiplied by the dilution factor of eleven to account for the reduced input and, therefore, make the stable *gpt*⁺-transfection frequencies from cotransfections more comparable with those from single plasmid transfections.

For the experiments whose results are given in table 3.4., selection for stable *gpt*⁺-transfectants, following the transfection experiment, was performed in normal, XHMPA selective-medium, and also in XHMPA selective-medium supplemented with potential transcription-inducing agents, as indicated. Where no stable transfectants were obtained following a transfection experiment, this is indicated by N.T.

Table 3.1. Comparative transfection-frequencies to *gpt*$^+$ in J558L for pPROM-Egpt constructs.

Plasmid	No. determinations and type of exp.		Transfection frequency (transfectants/cell x 10^4)	
			Range of frequencies	Average
pMMTV-Egpt	S:	1	9.0	9.0
	C:	4	9.9 → 13	12
pCMV-Egpt	S:	1	7.0	7.0
	C:	1	1.2	1.2
pMT-Egpt	S:	2	1.0 → 2.4	1.7
	C:	1	4.7	4.7
pML-Egpt	S:	1	1.6	1.6
	C:	3	1.2 → 3.6	2.0
pSV-Egpt	S:	0		
	C:	1	0.9	0.9
pIg-Egpt	S:	1	2.2	2.2
	C:	1	5.1	5.1
pTK-E(b')gpt	S:	1	1.0	1.0
	C:	1	6.5	6.5
pΔTK-Egpt	S:	7	0.8 → 8.3	4.0
pΔTK-E'gpt	C:	1	3.2	3.2

5'-direction resulted in an eight-fold decrease in expression of a downstream test-gene.

In spheroplast-fusion cotransfections (see section 4.2.3.) of the pPROM-Egpt constructs with another plasmid (pTK-E(b')ova, see section 4.2.2. and figure 4.1.), the resulting stable *gpt*$^+$-transfection frequencies dropped roughly in proportion to the decrease in the number of *gpt* sequences present in the transfection. This is consistent with the assumption (see section 3.2.) that the number of *gpt* genes introduced in the transfection experiment correlates with the subsequent average-level of XGPRT expression from the transfected population and, therefore, the stable *gpt*$^+$-transfection frequency. As the number of *gpt* genes present in cotransfections was about eleven-fold lower than in normal transfections, the transfection frequencies obtained from cotransfections, which are presented in the table 3.1., have been multiplied by this dilution factor, to make them more comparable with single-plasmid transfections.

3.4. pPROM-gpt constructs.

The IgH-gene enhancer sequence was deleted from each of the pPROM-Egpt constructs, to produce the corresponding pPROM-gpt construct (see figure 3.7.). Hence, these constructs contain only a promoter sequence as an expression-element upstream of the *gpt* coding-sequence. Whenever possible, a deletion of 1.2Kb was made between the *Eco* RI site at the 5'-end of the promoter fragment, and an *Acc* I site which is situated 0.2Kb into the pBR-derived sequence. This manipulation was carried out on the pPROM-Egpt constructs containing the ΔTK, IgH, MT-I, Adenovirus-ML and SV40 early promoters. The same deletion was not feasible for the MMTV-LTR and CMV-IE-promoter containing constructs. Instead, a deletion was made between the *Eco* RI and *Nde* I sites, which removed an additional 0.05Kb of pBR sequence. All deletions were made by cutting with the appropriate enzymes, filling-in with Klenow and then ligating the backbone. An additional pPROM-gpt construct, containing the Rous-sarcoma-virus(RSV)-LTR as a promoter element, was made by replacing the *Hind* III - *Bam* HI fragment in pRSV-CAT, which contains *cat*-SV40 sequences, with the 1.9Kb *gpt*-SV40 *Hind* III - *Bam* HI fragment from pSV$_2$gpt which is present in all other pPROM-gpt constructs. Hence, pRSV-gpt has the same non-promoter sequences as all

Figure 3.7. pPROM-gpt constructs.

The structures of the pPROM-gpt constructs, which contain only a promoter element to drive *gpt* transcription, are illustrated. The plasmids differ in the promoter element upstream of the *gpt* sequence, and there are also small differences in some of the pBR backbone sequences immediately upstream of the promoter elements. All the constructs (except for pRSV-gpt) were made by deleting the IgH enhancer from the corresponding pPROM-Egpt construct (see section 3.4.). The constructs are labelled as follows:

1) pΔTK-gpt.
2) pMMTV-gpt.
3) pIg-gpt.
4) pMT-gpt.
5) pCMV-gpt.
6) pSV-gpt.
7) pML-gpt.
8) pRSV-gpt.

Thin lines represent the pBR backbones, and the position of the ampicillin-resistance gene (amp) is indicated. The *gpt* sequence is represented by a box filled with thick diagonal lines and SV40 fragments downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequences derived from the late region (SV40(L))). The promoter sequences are represented by boxes filled with the following patterns: ΔTK (unfilled), MMTV-LTR (triangles), IgH (dots), MT-I (broken zig-zags), CMV-IE (thin diagonal lines), SV40 early (squares), Adenovirus major-late (ML) (broken horizontal lines), RSV-LTR (vertical lines).

The plasmids are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Restriction sites are abbreviated as follows: *Acc* I: A, *Eco* RI: R, *Hind* III: H, *Bam* HI: B, *Nde* I: N. Only a partial restriction map is shown, with emphasis on the sites where the enhancer was deleted during construction of the plasmids (i.e. upstream of the promoter). Sites deleted during construction of the plasmids are indicated with asterisks.

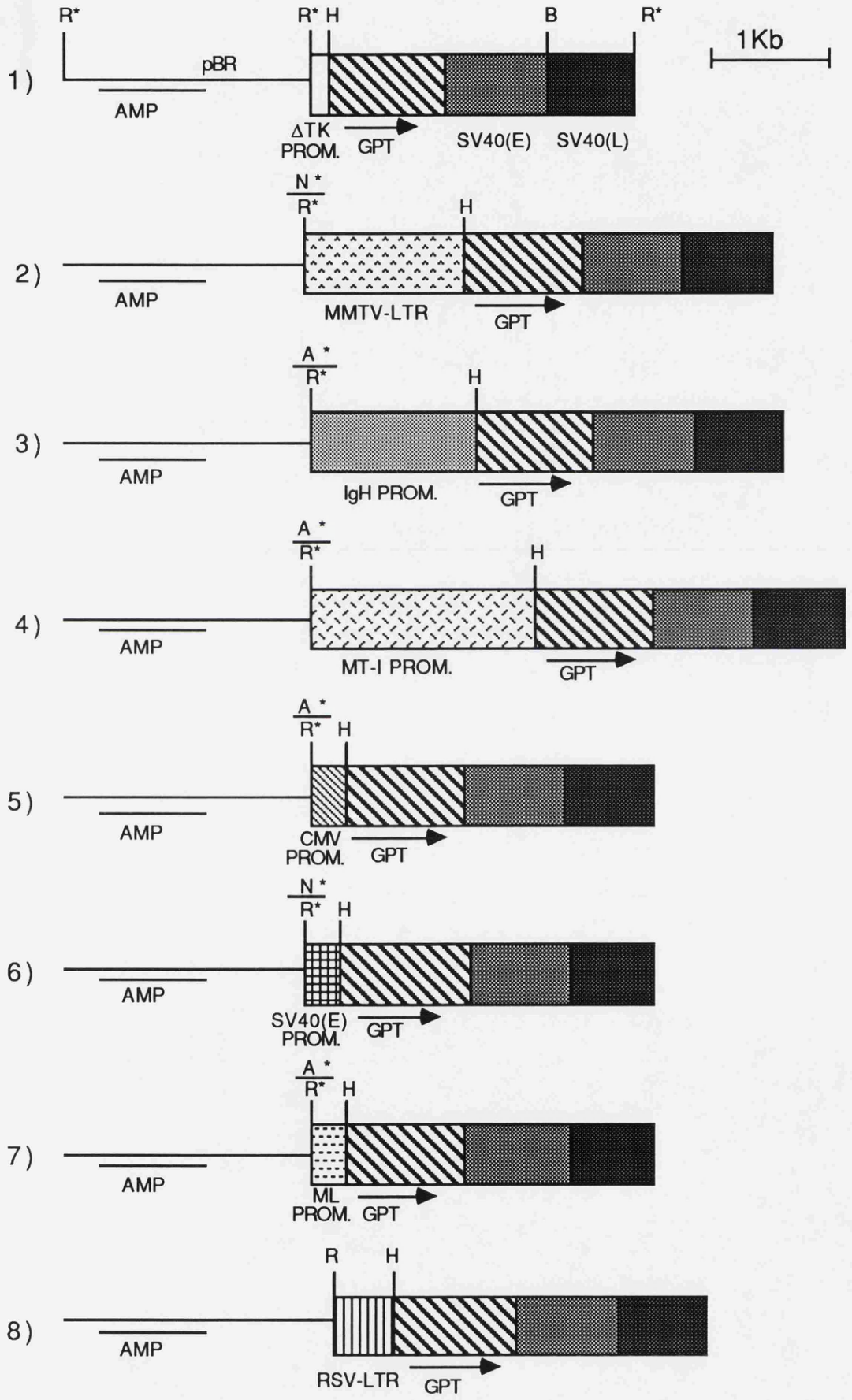

1)
R*
R* H
B
R*
1Kb
pBR
AMP
ΔTK
PROM.
GPT
SV40(E)
SV40(L)
2)
N *
R*
H
AMP
MMTV-LTR
GPT
3)
A *
R*
H
AMP
IgH PROM.
GPT
4)
A *
R*
H
AMP
MT-I PROM.
GPT
5)
A *
R* H
AMP
CMV
PROM.
GPT
6)
N *
R* H
AMP
SV40(E)
PROM.
GPT
7)
A *
R* H
AMP
ML
PROM.
GPT
8)
R
H
AMP
RSV-LTR
GPT

Table 3.2. Comparative transfection-frequencies to *gpt*$^{+}$ in J558L for pPROM-gpt constructs.

Plasmid	No. determinations and type of exp.		Transfection frequency (transfectants/cell x 10^4)	
			Range of frequencies	Average
pMMTV-gpt	S:	2	0.33 → 0.34	0.34
pMT-gpt	S:	1	0.14	0.14
	C:	1	0.73	0.73
pCMV-gpt	S:	1	1.10	1.10
pIg-gpt	S:	1	0.04	0.04
pML-gpt	S:	2	N.T. → 0.01	0.01
pRSV-gpt	S:	3	N.T. → 0.02	0.01
pΔTK-gpt	S:	2	N.T. (<< 0.002)	< 0.002
pSV-gpt	S:	1	N.T. (<< 0.002)	< 0.002

N.B. See legend associated with table 3.1.

other pPROM-gpt constructs, except it contains an additional 0.2Kb or 0.25Kb of pBR sequence upstream of the RSV-LTR.

Table 3.2. shows that low, stable *gpt*$^+$-transfection frequencies were generally obtained, and in some cases no stable transfectants were obtained at all, following introduction of pPROM-gpt constructs into J558L. The comparative stable-transfection frequencies with pPROM-Egpt and pPROM-gpt constructs (see tables 3.1. and 3.2.) demonstrate the overwhelming importance of the IgH enhancer in ensuring efficient XGPRT-expression from the *gpt* sequence in J558L. When the IgH enhancer was placed upstream of the promoter (pPROM-Egpt constructs), transfection frequencies were always increased, and in most cases by at least an order of magnitude (the only exception was with constructs containing the CMV-IE promoter, where the increase was only four-fold, but this result was based on single transfections).

It might be argued that the poor transfection-frequencies generated by pPROM-gpt constructs relative to their corresponding pPROM-Egpt constructs, are due to the reduction in size of the plasmids (1.2 - 1.3Kb deletion) having some unanticipated physical effect, such as a change in the conformational characteristics of the plasmids. Therefore, I made a deletion between the *Pvu* II and *Acc* I sites of pΔTK-Egpt, by digesting with *Acc* I (and filling-in) and *Pvu* II and then ligating the backbone, to produce pΔTK-E'gpt (see figure 3.6.). This removed the same sequences which were removed in construction of pPROM-gpt constructs, except for a 0.3Kb *Pvu* II - *Eco* RI fragment, which contains most of the recognised elements of the IgH enhancer (Banerji *et al.*; Gillies *et al.*; Neuberger, all 1983). This deletion might be expected to cause a similar change in the plasmid structure to the possible change when a pPROM-gpt construct is made from a pPROM-Egpt construct. However, pΔTK-E'gpt gave a stable *gpt*$^+$-transfection-frequency as high as pΔTK-Egpt (see table 3.1.). This further supports the view that the presence of the enhancer is the crucial factor in determining the expression level from *gpt*.

3.5. gpt containing plasmids without a promoter.

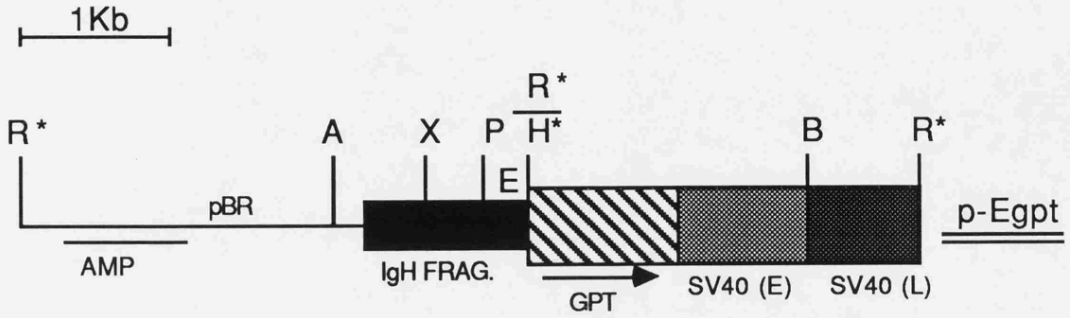
p-Egpt, a construct containing only the enhancer upstream of the *gpt* coding-sequence, was made by deleting the ΔTK promoter of pΔTK-Egpt

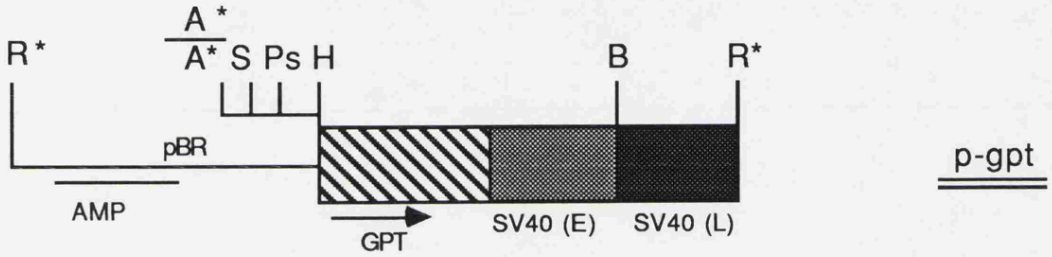
Figure 3.8. Promoterless, gpt expression-plasmids.

The constructs illustrated both contain a *gpt* coding-sequence and SV40 sequences downstream, but do not contain a recognised promoter element upstream of the *gpt* sequence. The only expression element upstream of the *gpt* sequence in p-Egpt is the IgH-gene enhancer, and p-gpt contains no upstream expression-element at all. The construction of these plasmids is described in section 3.5.

Thin lines represent the pBR backbone, and the position of the ampicillin-resistance gene (amp) is indicated. The IgH gene major-intron fragment, containing the enhancer (E), is represented by a thin black box, the *gpt* sequence is represented by a box filled with thick diagonal lines and SV40 sequences are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the fragment containing sequences derived from the late region (SV40(L))).

The plasmids are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Restriction sites are abbreviated as follows: *Acc* I: A, *Xba* I: X, *Pvu* II: P, *Eco* RI: R, *Hind* III: H, *Bam* HI: B, *Sal* I: S, *Pst* I: Ps. Only partial restriction maps are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.


1Kb
R*
A
X
P
R*
H*
E
B
R*
pBR
AMP
IgH FRAG.
GPT
SV40 (E)
SV40 (L)
p-Egpt


A*
R*
A*
S
Ps
H
B
R*
pBR
AMP
GPT
SV40 (E)
SV40 (L)
p-gpt

(see figure 3.3.). This was done by cutting with *Eco* RI and *Hind* III, filling-in with Klenow and then ligating the backbone. p-gpt, a construct without any upstream expression elements at all, was made by digesting pCMV-Egpt with *Acc* I (in figure 3.6., the *Acc* I site in the polylinker sequence of pCMV-Egpt is represented as a *Sal* I site, which has the same recognition sequence), filling-in with Klenow (although not necessary for this manipulation) and then ligating the backbone. This resulted in a 1.5Kb deletion encompassing all CMV-IE - derived and IgH-gene - derived sequences, and also 0.2Kb of pBR sequence. The plasmids, p-Egpt and p-gpt, are illustrated in figure 3.8.

Interestingly, p-Egpt gave a stable *gpt*$^+$-transfection frequency almost as high as with pPROM-Egpt constructs, even though it does not contain a *bone fide* promoter element (see table 3.1. and 3.3.). It was proposed that because the enhancer is such an active element in J558L, it is able to promote use of cryptic TATA-boxes to drive *gpt* transcription. This would be similar to the way in which sterile μ-transcripts are driven off secondary or cryptic promoters in IgH genes, during early stages of B-lymphocyte development, prior to the recombination event that allows μ-transcripts to be driven from the promoter upstream of the V_H-region (Nelson *et al.* (1983), and see section 1.12.). Hence, the 5'-end points of *gpt* transcripts driven off this plasmid were estimated using primer-extension analysis, in order to identify the site of transcription initiation (see section 3.6.).

Comparison of the stable *gpt*$^+$-transfection frequencies produced by pTK-E(b')gpt and pΔTK-Egpt (see table 3.1.) also suggests that very limited sequences, in addition to the IgH enhancer, are required to exert a nearly optimal effect on gene expression. pΔTK-Egpt (see figure 3.3.) contains a truncated HSV-TK promoter (a 0.12Kb fragment downstream of the *Eco* RI site located within the promoter) which includes the TATA-box plus an additional, weak Sp1-binding site (GC hexanucleotide), but has lost an inverted CAAT-box and an upstream, strong Sp1-binding site, which are reported to play dominant roles in the activity of the promoter in *Xenopus* oocytes (see McKnight and Tjian, 1986). Nevertheless, this truncated element combined successfully with the IgH enhancer (essentially an enhancer and a TATA box, if the interactions of *trans*-acting factors with the promoter in J558L are the same as in *Xenopus* oocytes), to drive

Table 3.3. Comparative-transfection frequencies to *gpt*$^+$ in J558L for promoterless constructs.

Plasmid	No. determinations (all single plasmid transfections)	Transfection frequency (transfectants/cell x 10^4)	
		Range of frequencies	Average
p-Egpt	2	0.4 → 1.0	0.7
p-gpt	3	N.T. → 0.004	0.002

N.B. See legend associated with table 3.1.

expression from *gpt* as efficiently as in pTK-E(b')gpt (which contains the complete TK promoter; see figure 3.3.) and most pPROM-Egpt constructs (see table 3.1.).

One notable anomaly is that p-gpt, a plasmid without recognised expression-elements upstream of the *gpt* coding-sequence, gave rise to stable *gpt*$^+$-transfectants following two out of three transfection experiments, albeit at very low frequencies (see table 3.3.). This was possibly caused by integration of the plasmid into a site of the host chromatin where transcription of the *gpt* sequence(s) was initiated by endogenous transcription-signals. There are several reports in the literature which describe similar observations. For example, Hamamada (1986) demonstrated that an enhancerless gene which would normally not be expected to give transfectants, was activated into producing transfectants due to integration into a chromosomal-site adjacent to an active transcription-element, presumed to be an enhancer. Chromosomal position effects have also been observed for Retroviral integration into mouse embryos by Jaenisch *et al.* (1981), who showed that the chromosomal site for integration was often the determining factor for proviral expression. Other examples of position effects were described in section 1.9.1.3. If this is the reason why p-gpt was able to generate stable transfectants, it is difficult to explain why stable transfectants were not similarly generated following transfection with pΔTK-gpt and pSV-gpt (see table 3.2.). However, only one transfection was performed with each of these plasmids, and if further transfections were performed, stable transfectants might be similarly generated.

3.6. Primer extension of gpt transcripts in *gpt*$^+$ transfectants of J558L.

3.6.1. Strategy.

I wished to identify the 5'-end(s) of the *gpt* transcripts in p-Egpt transfectants in order to test the hypothesis made above (section 3.5.). Primer-extension analysis was used. This works on the principal that when a single-stranded DNA sequence is annealed to a complementary sequence

within a RNA transcript, the DNA sequence can act as a primer for the enzyme, reverse transcriptase, to synthesise a complementary DNA-strand in the 5'→3' direction to the 5'-end of the RNA transcript.

To allow detection and sizing of the extension products, the primer was radiolabelled. The primer was made using single-stranded M13mp9-5'gpt DNA as a template (see section 2.10.1.). M13mp9-5'gpt is a M13mp9 derivative which contains a 0.12Kb *Hind* III - *Bgl* II fragment, derived from the *gpt* gene (includes coding and 5'-non-coding sequences), inserted into the *Hind* III and *Sal* I (filled-in) sites of the polylinker sequence. The oligonucleotide primer, 5'-gpt (see section 2.10.1.2.), was annealed to the complementary *gpt*-sequence in single-stranded M13mp9-5'gpt DNA, and then extended along the M13mp9-5'gpt template in the presence of radiolabelled nucleotides using Klenow, to produce a radiolabelled DNA-strand, complementary to *gpt* RNA. This strand was released from the template by *Hind* III digestion and strand dissociation. Then it was isolated on a polyacrylamide gel (see section 2.10.1.3.) and recovered by electroelution (see section 2.7.7.2.1.). This primer was then annealed to poly(A)$^+$ RNA prepared from stable *gpt*$^+$-transfectants of J558L, and extended with reverse transcriptase (see section 2.12.4.).

Poly(A)$^+$-RNA preparations (see section 2.12.3.), isolated from stable *gpt*$^+$-transfectants containing one of four different *gpt* expression-plasmids, were analysed. The four constructs were pSV-Egpt, pΔTK-Egpt, pCMV-Egpt (see section 3.3. and figure 3.6.) and p-Egpt (see section 3.5. and figure 3.8.). The transcription-initiation sites of the SV40 early, complete HSV-TK and CMV-IE promoters, which are contained in the former three constructs, have been determined by other investigators (Reddy *et al.*, 1979; McKnight *et al.*, 1981; Akrigg *et al.*, 1985, respectively). It was expected that the same transcription-initiation sites would be used in the J558L transfectants and, therefore, the analysis of transfectants containing these promoters would serve as positive controls, indicating the accuracy with which the transcription-initiation site(s) could be identified. This was important because primer-extension analysis is prone to misleading interpretations, due to effects such as premature termination of the extension reaction.

The four constructs were expected to have the sequences shown below on the anti-sense (coding) strand around, and upstream of the

junction of the 3'-end of the expression element fragment with the 5'-end of the *gpt* sequence. The *Hind* III sites (or remains of the *Hind* III site in p-Egpt, because the site was deleted during construction of the plasmid; see section 3.5) at the junctions are underlined (_____). The previously identified transcription-initiation sites for each promoter are marked with asterisks (*), and the TATA (or Hogness) boxes are marked by a string of filled circles (●).

pSV-Egpt.

The *Hind* III site at the 3'-end of the SV40 early-promoter was ligated to the *Hind* III site at the 5'-end of the *gpt* sequence during construction of pSV-Egpt (see section 3.3.1. and figures 3.5. and 3.6.).

```
5'     ●●●●●●                        *   *
TTTTTTTATT TATGCAGAGG CCGAGGCCGC CTCGGCCTCT GAGCTATTCC AGAAGTAGTG
AGGAGGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAAGCTT → gpt.
                                   _____        3'
```

pΔTK-Egpt.

Krieg *et al.* (1984) ligated a *Hind* III linker, to a filled-in *Bgl* II site which defined the 3'-end of the HSV-TK-promoter fragment (see McKnight *et al.*, 1981). This *Hind* III site was ligated to the *Hind* III site at the 5'-end of the *gpt* sequence during construction of pΔTK-Egpt (see section 3.3.1. and figure 3.3.).

```
5'
●●●●●●                        * *
TATTAAGGTG ACGCGTGTGG CCTCGAACAC CGAGCGACCC TGCAGCGACC CGCTTAACAG
CGTCAACAGC GTGCCGCAGA TCCCAAGCTT → gpt.
                                   _____        3'
```

pCMV-Egpt.

In its native sequence, the human CMV-IE promoter fragment used extended to a *Sst* II site downstream of the transcription-initiation site (Akrigg *et al.*, 1985). This site was filled-in and ligated to the *Sma* I site of a polylinker sequence, to leave a polylinker sequence containing *Bam* HI, *Xba* I, *Sal* I, *Pst* I, *Hind* III sites downstream of the 3'-end of the

promoter fragment. The promoter was then inserted into another plasmid as a 0.3Kb *Eco* RI - *Hind* III fragment to generate the pUC7-CMV cloning plasmid (A. Akrigg, personal communication). This same *Eco* RI - *Hind* III fragment was inserted upstream of the *gpt* sequence during construction of pCMV-Egpt (see section 3.3.1. and figure 3.6.).

```
5'   ●●●●● ●●                          *
AGGTCTATAT AAGCAGAGCT CGTTTAGTGA ACCGTCAGAT CGCCTGGAGA CGCCATCCAC
GCTGTTTTGA CCTCCATAGA AGACACCGGG ACCGATCCAG CCTCCGGGGA TCCTCTAGAG
TCGACCTGCA GCCCAAGCTT → gpt.
                  ______        3'
```

<u>p-Egpt</u>.

The ΔTK promoter was deleted from pΔTK-Egpt, by cutting with *Eco* RI and *Hind* III, and then the ends of the backbone were filled-in and ligated, to produce p-Egpt (see section 3.5. and figure 3.7.). This left a 0.7Kb IgH-gene fragment (Banerji *et al.*; Gillies *et al.*, 1983), which contains the enhancer sequence, directly upstream of the *gpt* sequence. Strings of triangles (▲) marked above or below the IgH-gene-derived sequence emphasise where the sequence is AT-rich, and a filled rectangle (▮) emphasises an A-residue six bases upstream of the fragment containing the *gpt* sequence; these are of relevance to the discussion in section 3.6.2.

```
5'                      ▲▲▲▲▲▲▲▲▲▲ ▲▲
CCGAAACTGG AGAGGTCCTC TTTTAACTTA TTGAGTTCAA CCTTTTAATT TTAGCTTGAG
TAGTTCTAGT TTCCCCAAAC TTAAGTTTAT CGACTTCTAA AATGTATTTA GAATTAGCTT → gpt
                 ▲▲▲▲▲▲▲▲▲▲▲▲▲▲                    ▮      _____      3'
```

3.6.2 Results and discussion.

Before making the radiolabelled *gpt*-RNA - complementary primer, it was necessary to ensure that the 5'-gpt 17mer-oligonucleotide primer anneals specifically to the complementary *gpt*-sequence in the single-stranded M13mp9-5'gpt DNA template. Therefore, two sequencing reactions (see section 2.10.2.) were performed on M13mp9-5'gpt, one using the M13

universal-primer (which is complementary to the M13 sequence immediately 3' of the *gpt* insert; see figure 2.1.) and the other using the 5'-gpt primer (which is complementary to part of the *gpt* insert; see figure 2.1. and 3.10.). The sequences of these primers are shown in section 2.10.1.2. In both cases readable sequences were produced (not shown), which confirms that the 5'-gpt - oligonucleotide primer, as well as the universal primer, anneal to predominantly one site on the template. Also, the two sets of sequencing reactions were run alongside each other, and on the resulting autoradiogram the first readable bases from the two sequencing reactions were, within experimental limits (a difference of one base), the same distance from the 3'-end of their respective oligonucleotide primer. This assumes that the primers anneal to the M13mp9-5'gpt template at the expected locations. The site of annealing of the universal primer is, of course, well characterised and, hence, this strongly implies that that the 5'-gpt primer anneals, as anticipated, to the complementary *gpt* sequence in the M13mp9-5'gpt DNA-template.

The radiolabelled primer complementary to *gpt* mRNA was then made and the extension reactions performed. The extension reaction products were run-out on sequencing gels. Sequencing-reaction tracks of M13mp9-5'gpt, primed with 5'-gpt - oligonucleotide, were run-out alongside (see figure 3.9.). The known size of comigrating fragments in the sequencing-reaction tracks allows the primer-extension products to be sized (see figure 3.10.). The primer was present in large excess for all reactions, as indicated by the intense non-extended primer band on the autoradiograms of figure 3.9. The primer was 69 bases long, instead of the expected 65 bases, which is the distance from the 5'-end of the 5'-gpt primer to the site of *Hind* III cleavage at which the *gpt*-RNA - complementary radiolabelled primer was detached from the radiolabelled M13-backbone strand, after extending from the 3'-end of the 5'-gpt primer along the M13mp9-5'gpt template. Presumably, following *Hind* III cleavage of the Klenow-extension-reaction product, but before strand separation (see section 2.10.1.3.), residual Klenow activity filled-in the 5'-overhang to extend the radiolabelled strand by four bases.

The estimated major extension-lengths (from the autoradiograms shown in figure 3.9.) from the *gpt*-RNA - complementary primer are 57 and 59 bases for ΔTK-promoter - generated transcripts, 58 and 62 bases for the

Figure 3.9. Primer-extension analysis of gpt transcripts in stable transfectants.

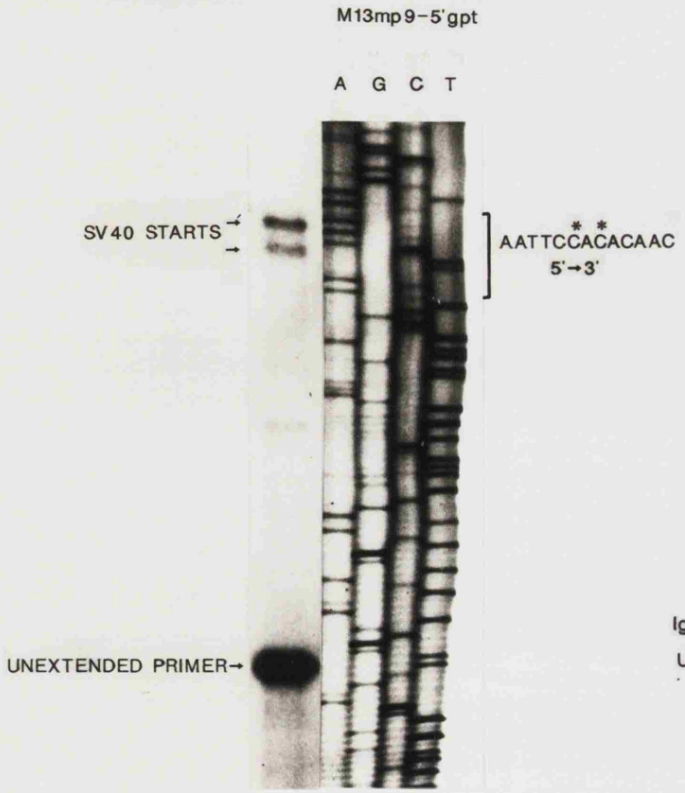
A radiolabelled single-stranded DNA primer, complementary to *gpt* RNA, was prepared using M13mp9-5'gpt (see section 2.10.1.1.) single-stranded DNA template and the 5'-gpt oligonucleotide-primer (see section 2.10.1.2.) using the method described in sections 2.10.1.3. and 3.6.1. It was then hybridised to several poly(A)$^+$-RNA preparations, from stable *gpt*$^+$-transfectants of J558L containing different *gpt*-expression plasmids. The primer was extended with AMV reverse-transcriptase, to what is assumed to be the 5'-end(s) of the *gpt*-mRNA template(s). Therefore, this indicates the initiation sites for *gpt* transcription in each *gpt* expression-plasmid. Primer-extension products were run on DNA-sequencing gels, alongside a sequencing reaction of M13mp9-5'gpt primed with the 5'-gpt oligonucleotide, as illustrated. Separate photographs of the primer-extension-reaction tracks and the sequencing-reaction tracks, from the same autoradiograms, are placed alongside each other. This is because the sequencing tracks are overexposed at exposures required to visualise the low-intensity signals from the primer-extension-reaction products. However, as the photographs were taken from the same position, the relative positioning of the bands is unaffected.

Autoradiogram A: a primer-extension products with poly(A)$^+$ RNA from a pSV-Egpt transfectant of J558L (clone B2). For this reaction, 1µg poly(A)$^+$ RNA was annealed to 5,000 cpms of the primer.

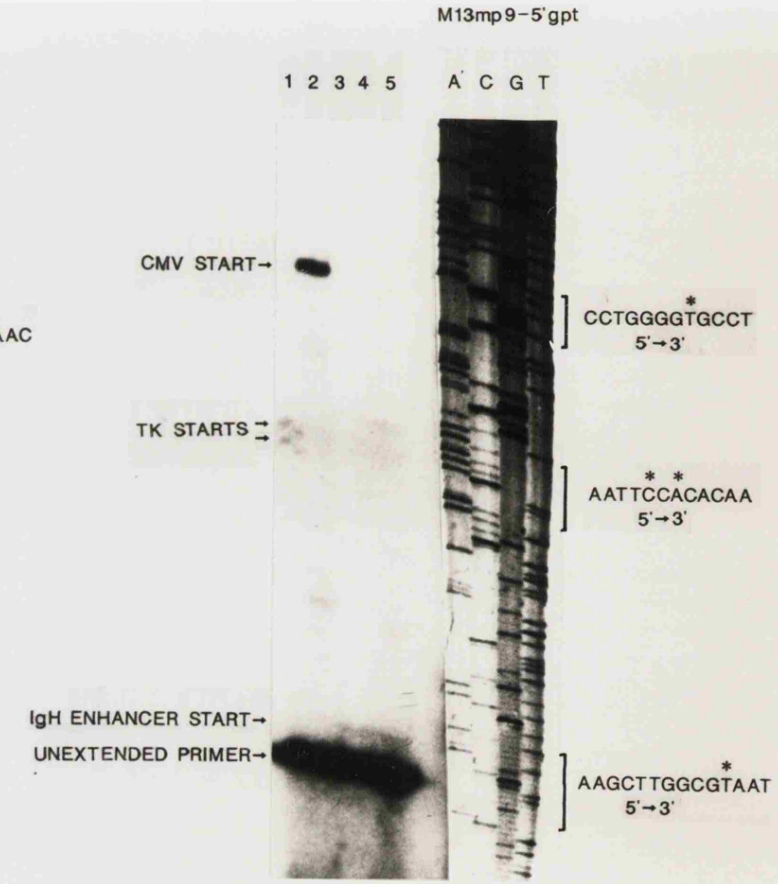
Autoradiogram B: primer-extension products following annealing of 5µg poly(A)$^+$ RNA to 5,000 cpms of the primer. The tracks contain the extension products of RNA, isolated from stable *gpt*$^+$-transfectants of J558L containing the following plasmids: track1: pΔTK-Egpt (clone B4); track 2: pCMV-Egpt (clone F2); track 3: p-Egpt (polyclonal transfectant-population F3); track 4: p-Egpt (clone D3); track 5: p-Egpt (clone D4).

The unextended, *gpt*-RNA - complementary primer comigrates with a fragment in the M13mp9-5'gpt sequencing tracks, which has a T residue at its 3'-terminus. The primer and sequencing reaction fragments have the same 5'-end, because they were made using the same oligonucleotide primer (5'-gpt). Hence, the primer sequence extends as far as the penultimate A residue of the *Hind* III site, at the 5'-end of the *gpt* sequence in the M13mp9-5'gpt template strand (see figure 3.10.). This means that the primer is four nucleotides longer than the product that is expected following *Hind* III cleavage of the Klenow-extension-reaction product (see section 2.10.1.3. for details of making probe). This is presumably because residual Klenow activity filled-in the 5'-overhang, following *Hind* III digestion, but before the strand-dissociation step.

Ideally, a primer-extension reaction on RNA from untransfected J558L is performed, as a negative control. However, the different tracks control each other and, therefore, bands specific to a particular transfectant are easily identified. Non-specific bands, which appear in all tracks, are ignored. Alongside the major products of the primer-extension reactions, the sequence read from the comigrating M13mp9-5'gpt sequencing-reaction fragments is read. The nucleotides, at the 3'-termini of sequencing-reaction fragments which are estimated to comigrate with the primer-extension products, are indicated with asterisks. The sequence of this region of M13mp9-5'gpt (the non-coding sequence, instead of the coding sequence which is read from the sequencing reaction) is given in full in figure 3.10., and the results are analysed in detail in section 3.6.2.


M13mp9-5'gpt
A G C T
SV40 STARTS
AATTCCACACAAC
5'→3'
UNEXTENDED PRIMER→

A


M13mp9-5'gpt
1 2 3 4 5
A C G T
CMV START→
CCTGGGGTGCCT
5'→3'
TK STARTS
AATTCCACACAA
5'→3'
IgH ENHANCER START→
UNEXTENDED PRIMER→
AAGCTTGGCGTAAT
5'→3'

B

Figure 3.10. Comigrating gpt-transcript primer-extension products and M13mp9-5'gpt sequencing-reaction fragments.

Part of the M13mp9-5'gpt anti-sense strand (the sequence complementary to that read from the sequencing tracks in figure 3.9.), including both the 5'-end of the *gpt* sequence and upstream, M13 sequences, is shown. The M13 sequence, from the autoradiogram in figure 3.9., appears to agree with that of Yanisch-Perron *et al.* (1985), and the *gpt* sequence (not all visible on the sequencing gels) is taken from the sequence published by Pratt and Subramani (1983).

The sizes of the primer-extension products of *gpt* transcripts, from the different *gpt* expression-plasmids, are indicated by symbols which show the end-points of fragments in the sequencing reactions which comigrated, on the gels of figure 3.9., with the primer-extension products. The symbols signify the following results:

* - the terminal nucleotide of the fragment from the M13mp9-5'gpt sequencing-reaction, with which the major extension product from the pCMV-Egpt transfectant comigrated.
⊙ - the terminal nucleotide of the fragment from the M13mp9-5'gpt sequencing-reaction, with which the major extension products from the pSV-Egpt transfectant comigrated.
● - the terminal nucleotide of the fragment from the M13mp9-5'gpt sequencing-reaction, with which the major extension products from the pΔTK-Egpt transfectant comigrated.
▮ - the terminal nucleotide of the fragment from the M13mp9-5'gpt sequencing-reaction, with which the major extension product from p-Egpt transfectants comigrated.

≡≡≡≡≡≡ - sequence complementary to the 5'-gpt oligonucleotide primer.

Lines (_____) are drawn above *Hind* III (AAGCTT) and *Bgl* II (AGATCT) sites. These sites define the junction of the 0.12Kb 5'-*gpt* fragment with the M13mp9 polylinker sequence.

```
5'                    *                                                ⊙
--M13-- AC TCATTAGGCA  CCCCAGGCTT  TACACTTTAT  GCTTCCGGCT  CGTATGTTGT

 ●⊙●                                                          ▮
GTGGAATTGT GAGCGGATAA  CAATTTCACA  CAGGAAACAG  CTATGACCAT  GATTACGCCA

_____gpt →                                                 3'-≡≡≡≡≡≡≡≡
AGCTTGGACA CAAGACAGGC  TTGCGAGATA  TGTTTGAGAA  TACCACTTTA  TCCCGCGTCA

≡≡≡≡≡≡≡≡≡-5'
GGGAGAGGCA GTGCGTAAAA  AGACGCGGAC  TCATGTGAAA  TACTGGTTTT  TAGTGCGCCA

_____
GATCTCGACG GATCCCCGGG  AATTCACTGG  CCGTCGTTTT  ACAACGTCGT  -gpt -- M13
                                                                  3'
```

SV40 early-promoter - generated transcripts and 101 bases for the CMV-IE promoter - generated transcripts (see figure 3.10.). These are only estimates, because defining which fragments in the sequencing tracks comigrate with the primer-extension-reaction products is rather subjective, particularly for the largest primer-extension-reaction fragments where the comigrating sequencing-reaction fragments are closely spaced. However, the above estimates agree with the previously reported transcription-initiation sites for each promoter. This is the expected result, because many previous studies have shown that enhancers do not alter the transcription-initiation sites in promoters they stimulate (see for example, Banerji *et al.*, 1981; Levinson *et al.*, 1982; Banerji *et al.*, 1983). This result further supports the view that the differences in transfection frequencies between pPROM-Egpt and pPROM-gpt constructs (see section 3.4., and tables 3.1.and 3.2.) are due to the influence of the IgH enhancer on the rate of transcription initiation of the *gpt* sequence. This is because the above result suggests that pPROM-Egpt and pPROM-gpt constructs, containing the same promoter element, utilise the same transcription-initiation site and, therefore, produce identical transcripts. If so, the differences in transfection frequencies cannot be due to post-transcriptional influences (see section 1.9.). More importantly, these results seem to illustrate the accuracy with which the primer-extension technique identified transcription-initiation sites. Therefore, this suggests the initiation site(s) determined for *gpt* transcripts driven off p-Egpt is also likely to be accurate.

The primer-extension products from RNA of the p-Egpt transfectants (two monoclonal populations and one polyclonal population) were the most difficult to detect. All but one of the transfectants analysed were monoclonal and, therefore, the levels of *gpt*-RNA expression do not have a high statistical significance. Even so, it is worth pointing out that the relatively low level of *gpt* RNA in all three p-Egpt transfectant populations is consistent with the observation that stable gpt^+-transfection frequencies obtained with p-Egpt are lower than with pPROM-Egpt constructs (see tables 3.1. and 3.3.).

As seen in figure 3.9., although the extension-product signals are weak, the same major, specific extension-product was produced from RNA derived from the polyclonal and two monoclonal p-Egpt transfectants. The primer appears to be extended by 6 bases to produce this major extension-

product, which implies that the predominant transcription-initiation site is only 6 bases upstream of the remains of the deleted *Hind* III recognition-site at the 5'-end of the *gpt* sequence. Analysis of the Ig-sequence upstream of the *Hind* III site in p-Egpt (see coding-strand sequence upstream of the *gpt* sequence, which is shown in section 3.6.1.) suggests that transcription is, therefore, initiated at an A residue in the coding strand; this is the base most favoured at transcription-initiation sites by RNA-polymerase II (Breathnach and Chambon, 1981). Several repeat sequencing gels are required, however, to precisely define the transcription-initiation site. Also, a sequencing reaction of p-Egpt sequences in this region is required to ensure that the fusion of filled-in *Eco* RI and *Hind* III sites, used in the construction of p-Egpt (see section 3.5.), produced the anticipated sequence shown in section 3.6.1., on which the conclusion that transcription started at an A residue is based. Therefore, a more significant observation, less likely to be affected by imperfect fusion of the sites during construction of p-Egpt, is the presence of an AT-rich sequence stretching from 20 to 33 bases upstream of the identified major transcription-initiation site. Breathnach and Chambon (1981) and Butcher and Trifonov (1986) described a consensus TATA (Hogness) box for eucaryotic promoters as TATAAA, which starts about 30 bases upstream of the initiation site. Therefore, it seems that this IgH-gene - derived sequence acts as a TATA box, even though the colinearity of specific bases with the consensus TATA-box is weak. Hence, this TATA box, in combination with the downstream A-residue, and under the influence of the upstream IgH-enhancer elements, seems to act as a cryptic promoter. There appears to be at least one other minor extension-product from RNA of p-Egpt transfectants (barely visible in figure 3.9.), in which the primer was extended by approximately 64 bases. This also corresponds to a transcription-initiation site, on the coding strand in p-Egpt, which is about 30 bases downstream of an AT-rich region (see sequence in section 3.6.1.). Hence, this sequence may also act as part of a cryptic promoter.

3.7. Attempt to induce expression-elements.

3.7.1. Background and strategy.

Both the mouse MT-I promoter and the mouse MTV-LTR can be activated by specific inducing-agents. Mayo *et al.* (1982) demonstrated that the MT-I promoter is activated in HeLa cells and LTK$^-$ cells, by addition of cadmium, and also a range of other less efficient inducing agents such as glucocorticoids. Carter *et al.* (1984) identified the regulatory sequences within the MT-I promoter, when demonstrating an induction response to cadmium (and other metals) in monkey kidney (CV1) cells. Lee *et al.* (1981) showed that MMTV-LTR - containing vectors are subject to induction by glucocorticoids such as dexamethasone, in CHO cells, and the responsive sequences, in rat XCtk$^-$ cells, were mapped to the LTR region by Chandler *et al.* and Payvar *et al.* (1983). Some of the most studied *trans*-acting factors in higher eucaryotes are the steroid-receptor proteins, which are thought to bind to and activate steroid inducible *cis*-acting sequences following binding of the steroid (for reviews, see Anderson (1983) and Yamamoto (1985)).

In some cell lines, transcription levels from the MMTV-LTR and the MT-I promoter can be increased by more than an order of magnitude in the presence of the inducing agent. Hence, if a similar level of induction occurred on a *gpt* transcription-unit in J558L, I expected it to be detectable using the stable-*gpt*$^+$-transfection-frequency assay. It was not expected that these inducible elements would form the basis of a tightly controlled expression system, because, as seen in section 3.4., the elements show activity in the absence of induction; rather, their induction might be a means of obtaining yet higher levels of the products of transfected heterologous-genes. Before performing the transfections, a preliminary experiment was required to determine if the potential inducing agents, dexamethasone or $CdSO_4$, are toxic to J558L. The agarose-cloning efficiency was determined (see section 2.2.2.) in the presence of a range of concentrations of these inducing agents. Some of these concentrations severely impaired the cloning potential of the cells (see figure 3.11.).

J558L cells were then transfected with *gpt* expression-plasmids containing the MT-I promoter or the MMTV-LTR. Following the transfections, cell suspensions were split and *gpt* expression was selected for in the presence and in the absence of the appropriate potential inducing-agent. The concentration of inducing agent added was the highest possible non-toxic concentration. The cloning efficiency experiment described above

showed that 5μM dexamethasone (added from a 1mM solution in DMSO) or 1μM $CdSO_4$ can be added to the cells without affecting the cloning efficiency. These concentrations of inducing agents cause induction in other cell lines and, therefore, these concentrations were included in the selective medium following the transfections. Dexamethasone was added to cells following transfection with MMTV-LTR - containing plasmids, and $CdSO_4$ was added to cells following transfection with pMT-Egpt. The resultant, stable *gpt*$^+$-transfection frequencies were counted to see if they were greater than with non-induced transfectants.

3.7.2. Results and discussion.

The selection for transfectants was carried out in selective medium which gave particularly low transfection-frequencies. However, unlike in previous experiments, the interpretation of the results is not clouded by transfectional variation (see section 3.2.). This is because 'induction' and non-induction were applied to the same transfected cell population and the same selective medium was used (except for the presence or absence of inducer). Therefore, the cloning efficiency is incidental to the detection of an induction effect. The stable *gpt*$^+$-transfection frequencies obtained with pMT-Egpt and pMMTV-Egpt (see table 3.4.) do not suggest any induction effect by cadmium on the MT-I promoter, nor by dexamethasone on the MMTV-LTR. However, these pPROM-Egpt constructs are in themselves capable of producing high transfection-frequencies (see table 3.1.). It was proposed that under the influence of the IgH enhancer, the promoter elements are already in an 'induced' state and, therefore, the assay was for a 'super-induction' effect. Hence, I attempted to detect an induction effect on the pPROM-gpt construct, pMMTV-gpt, which does not contain the IgH enhancer and which gives lower stable-*gpt*$^+$-transfection frequencies than pMMTV-Egpt in the non-induced state (see tables 3.1. and 3.2.). However, again, there was no detectable induction effect by dexamethasone (see table 3.4.).

The most probable explanation for the lack of induction is that J558L does not contain the specific receptor molecule for the inducing agent, on which the induction effect is thought to be depandent. Druege *et al.* (1986) demonstrated that cell-types not responding to inducer

Figure 3.11. Cloning efficiency of J558L cells in the presence of potential transcription-inducing agents.

J558L cells were cloned in soft agarose (see section 2.2.2.), in the presence of several concentrations of the potential transcription-inducing agents, dexamethasone (added from a 1mM stock solution in DMSO) or $CdSO_4$, and also in the absence of any inducing agent, to act as a control. Cloning efficiencies at each concentration were determined in duplicate.

On the graph, the filled circles indicate the cloning efficiencies in the presence of $CdSO_4$, the filled squares indicate the cloning efficiencies in the presence of dexamethasone, and the broken line indicates the average cloning-efficiency in the absence of either agent.

It can be seen that at the higher concentrations of dexamethasone or $CdSO_4$, the cloning efficiency was severely reduced. This was an important finding, because if these concentrations of potential inducing-agents were added to cells following transfections, any induction effect might have been obscured by the toxic effect on the cells. The addition of 1μM $CdSO_4$ or 5μM dexamethasone did not affect the cloning efficiency and, therefore, these concentrations were included in the selective medium following transfection experiments (see section 3.7.).

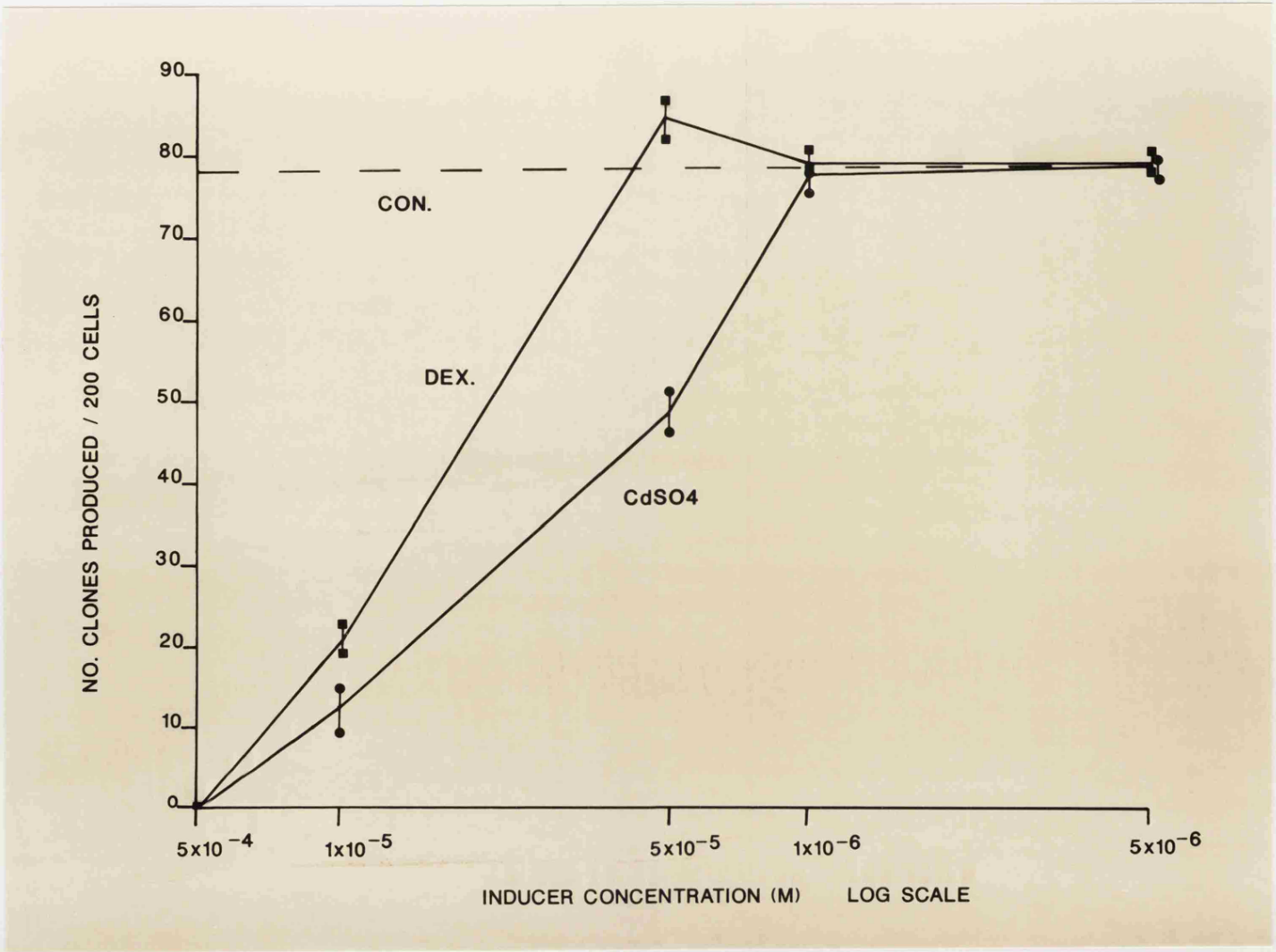

90
80
70
60
50
40
30
20
10
0
NO. CLONES PRODUCED / 200 CELLS
CON.
DEX.
CdSO4
5x10 −4
1x10−5
5x10−5
1x10−6
5x10−6
INDUCER CONCENTRATION (M) LOG SCALE

Table 3.4. Comparative transfection-frequencies to *gpt*$^+$ in J558L to test for activity of transcriptional inducers.

Plasmid	No. determinations and type of exp.		Transfection frequency (transfectants/cell x 10^4)	
			Range of frequencies	Average
pMMTV-Egpt	C: 2	(- DMSO/dex.)	0.70 → 0.83	0.77
	C: 2	(+ DMSO)	0.29 → 0.72	0.51
	C: 2	(+ DMSO/dex.)	0.44 → 0.74	0.59
pMMTV-gpt	S: 2	(-DMSO/dex.)	0.07 → 0.16	0.11
	S: 2	(+DMSO/dex.)	0.03 → 0.10	0.06
pMT-Egpt	C: 1	(- $CdSO_4$)	0.62	0.62
	C: 1	(+ $CdSO_4$)	0.96	0.96

N.B. See legend associated with table 3.1.

(estrogens), can become inducer responsive when cotransfected with a plasmid containing the inducible sequence and a plasmid expressing the inducer-receptor protein. A more trivial potential explanation for the absence of induction, is that the promoter elements in the transfected plasmids are not as responsive as the naturally-present promoter elements. This has been observed previously by Mayo *et al.* (1982), who showed that a MT-promoter - TK-coding-sequence hybrid-gene loses glucocorticoid inducibility following transfection into mouse cells.

3.8. Analysis of effect of Adenovirus translation-control elements in J558L.

3.8.1. Strategy and expression-plasmid construction.

As discussed in section 1.9.6., Kaufman (1985b) demonstrated that the translational efficiency of hybrid transcription-units, containing part of the non-coding tripartite-leader exon sequences of Adenovirus, can be increased in the presence of Adenovirus associated (VA) RNAs. It was decided to test if this effect could be reproduced on tripartite-leader-containing '*gpt*' mRNA in J558L, by using the stable-*gpt*$^+$-transfection-frequency assay. The construction of the *gpt* expression-plasmids, used in the assays to test for this effect, is described below and is illustrated in figure 3.12.

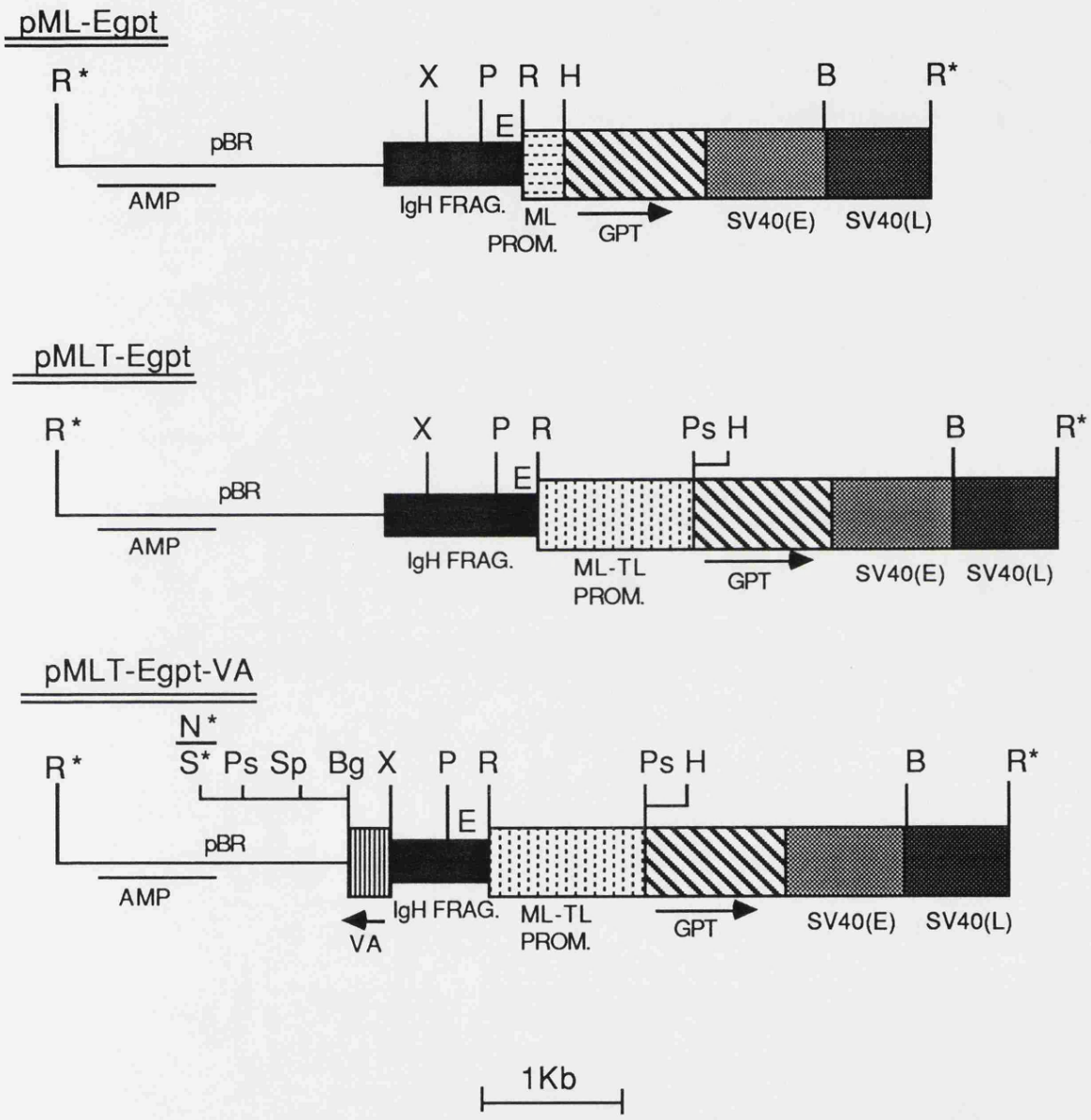
pD20 was a source of a 1Kb *Eco* RI - *Pst* I fragment. This fragment contains, reading from 5'→3', the Adenovirus ML promoter, a cDNA sequence of complete exons 1 and 2 and approximately 2/3 of exon 3 of the Adenovirus tripartite-leader (the *cis*-acting sequences involved in translational control) and a hybrid intron, consisting of a 5'-splice site from the third exon of the tripartite-leader and a 3'-splice site from a mouse Ig-gene V_H-region. A *Hind* III site was removed from pD20 by replacing the *Bam* HI - *Eco* RI backbone fragment with the *Bam* HI - *Eco* RI backbone from the plasmid, pXTKVδ6, which does not contain a *Hind* III site. Then the now unique *Hind* III site, in the intron sequence of the above described *Eco* RI - *Pst* I fragment, was deleted, by digesting with *Hind* III, filling-in with Klenow and then religating. The *Eco* RI - *Pst* I fragment was then cloned into the compatible sites of the pUC8 polylinker. It was then removed on an *Eco* RI -

Figure 3.12. Plasmid vectors to test for translation-control mediated through Adenovirus-derived elements.

The three plasmids, whose stable *gpt*$^{+}$-transfection frequencies were compared, to test if Adenovirus translation-control sequences could stimulate expression from *gpt* mRNA in J558L (see section 3.8.), are illustrated. pML-Egpt is a pPROM-Egpt construct previously described (see section 3.3.1. and figure 3.6.), and acted as the control. pMLT-Egpt differs from pML-Egpt in that the Adenovirus ML-promoter fragment (ML) is substituted by a ML-promoter fragment, which also contains sequences from the Adenovirus tripartite-leader and an an intron (ML-TL). pMLT-Egpt-VA is a derivative of pMLT-Egpt, in which the Adenovirus VA1-gene (VA) is substituted for a fragment upstream of the IgH enhancer.

pBR backbone sequences, containing the ampicillin-resistance gene (amp), are represented by thin lines, the IgH gene major-intron fragment, containing the enhancer (E), are represented by thin black boxes, the *gpt* sequence is represented by a box filled with thick diagonal lines, SV40 sequences are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequence derived from the late region (SV40(L))), the ML promoter is represented by a box filled with horizontal broken lines, the ML-promoter - tripartite-leader - intron fragment is represented by a box filled with vertical broken lines and the VA1-gene fragment is represented by a box filled with vertical lines.

The plasmids are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Restriction sites are abbreviated as follows: *Xba* I: X, *Pvu* II: P, *Eco* RI: R, *Hind* III: H, *Pst* I: Ps, *Nde* I: N, *Sph* I: Sp, *Bgl* II: Bg. For simplification, only partial restriction maps are shown, and emphasis is placed on the sites around the Adenovirus sequences. Sites delated during construction of the plasmids are indicated with asterisks.


pML-Egpt
R*
X
P
R
H
B
R*
E
pBR
AMP
IgH FRAG.
ML
PROM.
GPT
SV40(E)
SV40(L)
pMLT-Egpt
R*
X
P
R
Ps
H
B
R*
E
pBR
AMP
IgH FRAG.
ML-TL
PROM.
GPT
SV40(E)
SV40(L)
pMLT-Egpt-VA
N*
R*
S*
Ps
Sp
Bg
X
P
R
Ps
H
B
R*
E
pBR
AMP
IgH FRAG.
VA
ML-TL
PROM.
GPT
SV40(E)
SV40(L)
1Kb

Hind III cassette and cloned into pΔTK-Egpt, in the same way as for construction of all other pPROM-Egpt constructs (see section 3.3.1.), to produce pMLT-Egpt.

pMLT-Egpt was further modified by inserting a 0.3Kb *Sal* I (filled-in) - *Xba* I fragment (originally from Adenovirus 2, extending from the *Xba* I site at position 10579, to the *Bal* I site at position 10810; the *Bal* I site is linked to the polylinker sequence *Sph* I, *Pst* I, *Sal* I) obtained from the plasmid, pMHVA , into the *Nde* I (filled-in) - *Xba* I backbone. This resulted in the deletion of superfluous IgH-gene sequences and 0.25 Kb of pBR sequence, and replacement with the Adenovirus VAI-gene, to generate pMLT-Egpt-VA.

3.8.2. Results and discussion.

Stable *gpt*$^+$-transfection frequencies (see table 3.5.) do not indicate that the incorporation of Adenovirus translation-control signals into *gpt* expression-plasmids significantly increases the level of XGPRT expression. The transfection frequency was not significantly affected when the ML promoter in pML-Egpt, was replaced by the fragment containing the ML promoter, part of the Adenovirus tripartite-leader sequence (including the sequences in exons 2 and 3 which are reported to be involved in translational regulation) and an intron (pMLT-Egpt). pMLT-Egpt transfectants are expected to generate *gpt*-transcripts containing the tripartite-leader sequence in the 5'-non-coding region, but this was expected to be inconsequential because there is no VA RNA available in J558L to mediate translational stimulation of mRNA containing this sequence. The similar, stable *gpt*$^+$-transfection frequencies obtained with pML-Egpt and pMLT-Egpt suggests that the (assumed) difference at the 5'-end of the *gpt* transcripts has no effect on XGPRT expression-levels. However, the introduction of the VAI gene into pMLT-Egpt, also did not result in an increased transfection frequency with the resultant construct, pMLT-Egpt-VA. This was initially surprising considering VA RNA can, as discussed above, increase the translation rate of other hybrid transcription-units which contain the same tripartite-leader sequences (Kaufman, 1985b).

Although Kaufman's assays were performed on transiently infected cells and my assay was based on stably-transfected cells, this should not

Table 3.5. Comparative transfection-frequencies to *gpt*$^+$ in J558L with plasmids containing Adenovirus translational control elements.

Plasmid	No. determinations and type of exp.		Transfection frequency (transfectants/cell x 10^4)	
			Range of frequencies	Average
pML-Egpt	S:	1	1.6	1.6
	C:	3	1.2 → 3.6	2.0
pMLT-Egpt	S:	1	4.3	4.3
	C:	1	0.6	0.6
pMLT-Egpt-VA	S:	1	4.3	4.3
	C:	1	0.4	0.4

N.B. See legend associated with table 3.1.

be significant, because the effect is mediated at the translational level and is, therefore, independent of the site of DNA integration. I expected the VA gene to be cointegrated with the *gpt* gene at a high frequency, as was later demonstrated for another two-gene plasmid (see section 4.7.2.). Also, transcription of the VA gene is driven from a promoter which uses RNA-polymerase III, and so is likely to be constitutively active in most cell-types. However, evidence presented in chapter 5, suggests some kind of transcriptional interference effect may occur on plasmids containing two divergent transcription-units; the *gpt* and VA transcription-units in pMLT-Egpt-VA have this arrangement. Even so, I expected VA RNA to be produced in a reasonable proportion of pMLT-Egpt-VA transfectants. Kaufman (1985b) saw variability in the extent of translational stimulation in different cell-types, with the effect being particularly strong in COS-1 cells. It is possible that J558L may not be a suitable host-cell for this form of translational stimulation, perhaps due to the absence of an additional factor which is essential for the stimulation. In order to further the evidence that J558L is not an appropriate host, it would be necessary to show that *gpt* transcripts from pMLT-Egpt-VA contain the tripartite-leader sequences, and that the *gpt* mRNA is produced at the same levels as *gpt* mRNA from pML-Egpt (by RNA mapping and Northern blotting), and also to demonstrate that the VA RNA is present at a similar level to that in cells in which it activates translation (by Northern blotting).

3.9. Summary.

The relative activity of expression elements in the myeloma cell, J558L, was assessed by measuring their influence on expression of a test gene (*gpt*), using a stable-transfection-frequency assay. The results demonstrate the overwhelming influence of the IgH-gene enhancer on expression. The highest levels of expression are obtained when the IgH enhancer is combined with a promoter element to drive transcription, although the nature of the promoter only seems to be of secondary importance. Interestingly, the IgH enhancer can still drive expression in the absence of a promoter element, and this appears to be due to initiation of transcription from cryptic promoters within the IgH-gene sequence flanking the enhancer. Attempts to increase the expression level by

induction of the transcription rate, or by translational stimulation using Adenovirus elements, were unsuccessful.

CHAPTER 4: HIGH-LEVEL EXPRESSION FROM TRANSFECTED EUCARYOTIC-GENES IN J558L.

4.1. Introduction.

In chapter 3, I demonstrated that the highest level of XGPRT expression encoded by a transfected *gpt*-gene, is obtained when both the IgH-gene enhancer and a promoter element are present in the upstream flanking-region of the *gpt* sequence. The result most likely reflects a high efficiency of transcription initiation and, therefore, such a combination of *cis*-acting elements were expected to ensure efficient transcription of any downstream sequence. Two cDNA sequences, encoding chicken lysozyme (Land *et al.*, 1981) and chicken ovalbumin (McReynolds *et al.*, 1977) were available, and so these genes were used as models for obtaining expression of heterologous eucaryotic-proteins in J558L. Attempts were made to express these proteins by transfecting J558L with expression plasmids, which contained one of the cDNA sequences, flanked upstream by the IgH enhancer and a promoter.

Chicken lysozyme (1,4-β-N-acetylmuraminidase, EC 3.2.1.17.) is a small (mol.wt. = 14KD) monomeric protein. It is normally secreted and, therefore, the cDNA encodes a signal-sequence (Palmiter *et al.*, 1977) to enable delivery of the protein into the secretory pathway (see sections 1.9.7. and 1.10.1.). Lysozyme does not appear to undergo extensive post-translational modification, and can be expressed and secreted in a collection of lower-order expression hosts such as yeast (Oberto *et al.*, 1985) and *Xenopus* (Krieg *et al.*, 1984). Hence, it was expected that J558L would also possess the machinery necessary to allow lysozyme expression and secretion. If so, expression of lysozyme could be used to indicate the efficiency of an assembled transfection / expression system. Once an efficient system for lysozyme expression was developed, this could be used as the basis of attempts to assess the versatility of J558L as a host for the expression of more complex mammalian-proteins (see chapter 6).

Ovalbumin is another simple, monomeric protein (mol.wt. = 45KD) and is usually secreted. However, it differs from most other secreted proteins in not having a classic, N-terminal signal-sequence; instead an

internal signal-sequence located close to the N-terminus, which is not removed, ensures its delivery into the secretory pathway (Meck *et al.*, 1982). Even so, it was considered likely that this signal-sequence would be recognised in J558L, because other heterologous expression-hosts (e.g. *Xenopus* oocytes (Krieg *et al.*, 1984) and mouse L-cells (Sheares and Robbins, 1986)) can secrete ovalbumin .

One potential drawback in attempting to express these proteins is the lack of rapid, sensitive and simple assays for detecting their expression and activity. In fact ovalbumin is an inert protein with no assayable activity. An assay for lysozyme (Alderton *et al.*, 1945) involves determination of lytic activity on cultures of *Micrococcus lysodeikticus*. This was modified, for the detection of lysozyme, secreted from transformed yeast, by visualising lytic activity on bacterial lawns (Oberto and Davison, 1985). However, it was considered that this assay might not be sensitive enough to detect lysozyme at the concentration that might be secreted into the medium of a myeloma-cell culture (or that of any other mammalian expression-host). This would be particularly so during the early stages of development of an expression system, when the level of expression is unlikely to be optimal. Furthermore, the initial aim was to show that an expression system for heterologous proteins was functional, rather than to estimate the fidelity of the recombinant protein to the native protein by assaying for biological function. It was decided to screen for and quantify expression of the proteins from J558L transfectants by immune-precipitation (see section 2.5.2.), following biosynthetic-radiolabelling of cell protein (see section 2.5.1.). This approach did not require development of an assay system, as was required if a radioimmunoassay (RIA) or an enzyme-linked immunoassay (ELISA) assay was to be used.

Before measuring stable expression-levels from transfected heterologous genes, the small fraction of cells carrying the gene has to be identified and isolated. This entails the use of a selection system (see section 1.7.). As for most proteins, there is no selection-system available based on the expression of chicken ovalbumin or lysozyme. Therefore, a cotransfection of the non-selectable chicken gene with a selectable gene is necessary (see section 1.7.). As described in chapter 3, a *gpt* gene works effectively as a selectable marker for stable transfection in J558L.

Therefore, the non-selectable chicken cDNAs were cotransfected with a *gpt* gene.

4.2. Simple cotransfection.

4.2.1. Introduction and strategy.

Simple cotransfection is a term used, in this study, to describe cotransfections in which the selectable and non-selectable genes are located on separate expression plasmids (see section 1.7.).

Two ovalbumin expression-plasmids, pTK-E(b')ova and pMMTV-Eova, were made, as described in section 4.2.2., and as illustrated in figure 4.1. They are derivatives of the pPROM-Egpt constructs, pTK-E(b')gpt (which is not, strictly speaking, a pPROM-Egpt construct, but is referred to as such for simplification) and pMMTV-Egpt, respectively (see figure 3.6.). The *gpt* sequence was replaced with an ovalbumin-cDNA fragment, but the upstream IgH-enhancer and promoter sequences were maintained. As the parent constructs gave high, stable gpt^{+}-transfection frequencies (see table 3.1.), it was assumed that these upstream expression-elements can effectively initiate transcription and, therefore, these elements were expected to drive high-level transcription of the ovalbumin-cDNA sequence in the derivative plasmids, following introduction into J558L. The ovalbumin expression-plasmids were cotransfected with their parent pPROM-Egpt-construct, and in the case of pTK-E(b')ova, also with other pPROM-Egpt constructs. The only differences between the cotransfected plasmids are the coding sequences, the absence of the small t-antigen intron of SV40 in the ovalbumin expression-plasmids, and in some of the cotransfections with pTK-E(b')ova, different promoter elements. Homology between cotransfected plasmids is desirable, because cotransfected genes may cointegrate into the host genome at a higher frequency when they contain homologous sequences. This seems to be because homologous recombination between transfected plasmids plays at least a partial role in the generation of the concatameric units of transfected plasmids (Folger *et al.*, 1982). As described in section 1.6., the transfected plasmids may become integrated into the host genome in the form of a concatameric unit. Integration into the host genome usually occurs only at one site and, therefore, if

cointegration of the two plasmids is to occur, then this is likely to occur as part of a concatameric unit.

4.2.2. Construction of expression plasmids for simple cotransfection.

A *Hind* III ovalbumin-cDNA fragment was available within the plasmid, pTK_2ov (described in section 3.3.1.). To simplify manipulations, the *Hind* III site downstream of the cDNA was deleted, by partially cutting with *Hind* III, filling-in with Klenow and then religating, to produce pTK_2ov-Δ*Hind* III. This allowed a 1.6Kb *Hind* III - *Bam* HI fragment, containing the ovalbumin cDNA and the SV40, early polyadenylation-signal to be removed, and this was inserted into the *Hind* III - *Bam* HI backbone of pTK-E(b')gpt (see section 3.3.1.), to produce pTK-E(b')ova (see figure 4.1.).

As the MMTV-LTR was a highly-active element in the stable gpt^+-transfection frequency assays (see sections 3.3. and 3.4.), it was substituted for the TK promoter; this manipulation also positioned the enhancer closer to the LTR fragment than it was to the TK promoter in the parent construct. This manipulation involved digesting pTK-E(b')ova with *Xba* I and *Hind* III, which removed the TK promoter and most of the IgH-gene sequence including the entire enhancer sequence. Then a *Xba* I - *Hind* III fragment from pMMTV-Egpt (see section 3.3.1.), containing the IgH enhancer and the MMTV-LTR, was inserted into the backbone, to produce pMMTV-Eova (see figure 4.1.)

4.2.3. Simple cotransfection using spheroplast fusion.

pTK-E(b')ova was cotransfected with different pPROM-Egpt constructs into J558L by spheroplast fusion (see section 2.3.1.). The two plasmids were grown in separate bacterial hosts, and then the cultures were mixed with a ratio of pTK-E(b')ova carrying bacteria: pPROM-Egpt carrying bacteria (10:1 (v/v)). Spheroplasts were then made from this mixed suspension and used in the transfection.

These experiments were of use for two different reasons. Firstly, as discussed in section 3.3.2., they demonstrated that the number of *gpt* expression-plasmids added during a spheroplast-fusion transfection

Figure 4.1. Construction of ovalbumin expression-plasmids for use in simple cotransfection experiments.

The strategy used to construct two ovalbumin expression-plasmids, pTK-E(b')ova and pMMTV-Eova, which were used in simple cotransfections (see section 4.2.), is illustrated. The strategy is described in detail in section 4.2.2., but briefly, a 1.6Kb *Hind* III - *Bam* HI fragment, containing the ovalbumin-cDNA and the early polyadenylation-signal of SV40 downstream, was substituted for the *gpt* - SV40-early sequence in pTK-E(b')gpt, to generate pTK-E(b')ova. The IgH enhancer and TK promoter in pTK-E(b')ova, were then substituted by an IgH-enhancer - MMTV-LTR fragment, to generate pMMTV-Eova.

The TK promoter is represented by an unfilled box, the ovalbumin-cDNA is represented by a box filled with thin diagonal lines, the *gpt* sequence is represented by a box filled with thick diagonal lines, the MMTV-LTR is represented by a box filled with triangles, SV40 sequences downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragments containing the early polyadenylation-signal, with (SV40(E)) or without (SV40(E)pA) the small t-antigen intron, are shaded less darkly than the fragment derived from the late region (SV40(L))), the IgH gene major-intron fragments, containing the enhancer (E), are represented by thin black boxes and the pBR backbone, which contains the ampicillin- resistance gene (amp), is represented by a thin line.

The plasmids illustrated in full are shown as linearised at an *Eco* RI site for consistency with previous figures. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Pvu* II: P, *Hind* III: H, *Bam* HI: B, *Eco* RI: R, *Xba* I: X, *Sal* I: S, *Pst* I: Ps. Sites used in each manipulation performed are circled for emphasis. Only sites relevant to the manipulations are shown, for simplification. (P) indicated a partial digestion was performed. Sites deleted during construction of the plasmids are indicated with asterisks.

experiment correlates with the subsequent stable *gpt*$^{+}$-transfection frequency. Secondly, as described below, they also allowed an investigation of the frequency with which the ovalbumin cDNA is cointegrated, with the selected *gpt*-sequence, into the host genome.

Six stable *gpt*$^{+}$-transfectant clones were expanded. These were then analysed for cointegration of the ovalbumin and *gpt* sequences into the genomic DNA, by Southern-blotting analysis (see section 2.11.3.). Figure 4.2. (autoradiogram A) shows that in a Southern blot, containing *Hind* III / *Bam* HI digests of transfectant genomic-DNA, a *gpt*-SV40 single-stranded DNA probe (see section 2.10.1.) hybridised to fragments which comigrated with the 1.9Kb *Hind* III - *Bam* HI *gpt*-SV40 fragment of pSV$_2$gpt. This was the expected result because all pPROM-Egpt constructs contain this 1.9Kb *gpt*-SV40 sequence, and the fact that the cells were selected for stable expression from the *gpt* gene already suggested that *gpt* sequences were integrated. In contrast, in a blot containing *Hind* III digests of the same genomic DNA, which was hybridised to an ovalbumin-cDNA single-stranded DNA probe (see section 2.10.1.), no ovalbumin sequences could be detected in the genomic tracks (see figure 4.2., autoradiogram B), yet the signal intensity from an ovalbumin-cDNA positive-control fragment demonstrated the sensitivity of detection was similar to that for *gpt*-SV40 sequences. When this blot was stripped and then hybridised to the *gpt*-SV40 probe, hybridising sequences were again detectable in the genomic-DNA tracks (figure 4.2., autoradiogram C).

The control tracks on these blots did not contain carrier DNA. Hence, the hybridisation efficiency of DNA in such clean tracks may be different to that in tracks containing genomic DNA, where hybridising sequences are mixed with sequences non-complementary to the probe, which may interfere with the hybridisation process. Consequently, the copy number of the integrated sequences can not be determined accurately. Therefore, it can not be discounted that ovalbumin-cDNA sequences were present in the genomic DNA at a lower copy-number than *gpt* sequences.

Mulligan and Berg (1981a) showed that the number of *gpt* copies per transfected genome, following transfection with similar plasmids (pSV$_2$gpt and its derivatives), was small (range 1-5 copies). Therefore, and in view of the fact that the above experiment was a cotransfection, with a relatively low input of *gpt* expression-plasmids, it might be expected that

Figure 4.2. Southern blotting to analyse for cointegration of the ovalbumin cDNA with the gpt sequence following simple cotransfection by spheroplast fusion.

Genomic DNA was isolated from six monoclonal *gpt*$^+$-transfectants of J558L, which were generated by cotransfections with a pPROM-Egpt construct and pTK-E(b')ova, using spheroplast fusion (see section 4.2.3.). The genomic DNA was digested with *Hind* III, and with *Hind* III and *Bam* HI, run out on a 1% agarose gel, and then Southern blotted (see section 2.11.3.). The blots were hybridised to *gpt*-SV40 and/or ovalbumin radiolabelled single-stranded DNA probes (see section 2.10.1.), in order to estimate the cointegration frequency of the non-selected ovalbumin sequence with the selected *gpt*-sequence. Three autoradiograms are shown, and in each, the position and sizes of bands in tracks containing *Hind* III - digested λ DNA are indicated. Autoradiogram A shows a blot hybridised to the *gpt*-SV40 probe. A second blot, containing DNA from the same transfectants, was hybridised, first to an ovalbumin probe (autoradiogram B), and then stripped and hybridised to the *gpt*-SV40 probe (autoradiogram C).

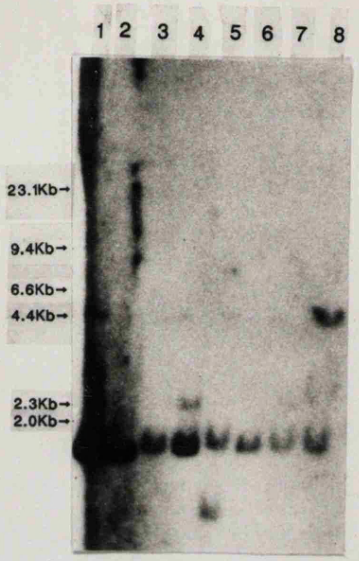
The tracks of the blot shown in autoradiogram A contained the following samples: tracks 1 and 2: *Hind* III / *Bam* HI digests of pSV_2gpt, containing 10pg and 1pg respectively of the 1.9Kb *gpt*-SV40 fragment (positive controls); tracks 3-8: 5μg of *Hind* III / *Bam* HI digests of genomic DNA from cotransfectants of pTK-E(b')ova and the following *gpt*-expression plasmids: track 3: pTK-E(b')gpt (clone C2); track 4: pTK-E(b')gpt (clone C9); track 5: pML-Egpt (clone E9); track 6: pML-Egpt (clone F2); track 7: pMT-Egpt (clone E7); track 8: pIg-Egpt (clone D5).

The tracks of the blot shown in autoradiograms B and C contained the following samples: tracks 1 and 2: *Hind* III digests of the plasmid, pSV_1ov, containing 20pg and 2pg respectively of the 1.4Kb ovalbumin-cDNA fragment (positive controls); tracks 3-8: 5μg of *Hind* III genomic-digests of DNA from cotransfectants of pTK-E(b')ova and the following *gpt* expression-plasmids: track 3: pTK-E(b')gpt (clone C2); track 4: pTK-E(b')gpt (clone C9); track 5: pMT-Egpt (clone E7); track 6: pIg-Egpt (clone D5); track 7: pML-Egpt (clone F2); track 8: pML-Egpt (clone E9).

Autoradiogram A shows that the *gpt*-SV40 probe hybridises to the 1.9Kb fragment present in pSV_2gpt, and also to the same fragment in the genomic-DNA tracks. This confirms that the selected *gpt*-sequence integrates into the genomic DNA.

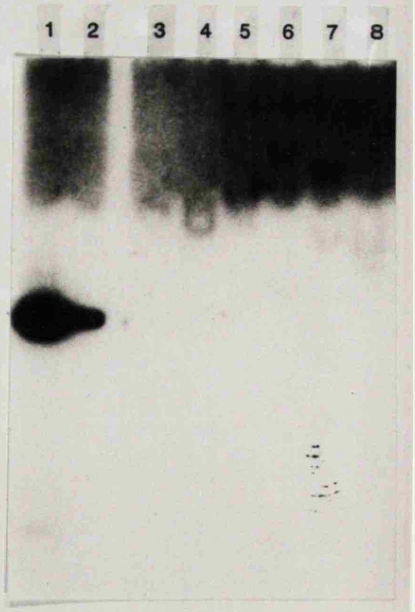
Ovalbumin sequences can not be detected in the genomic DNA of the same transfectants (autoradiogram B), even though, at the long exposure shown, the signal intensity from the control track is higher than that from the control tracks in autoradiogram A. Autoradiogram C shows, despite high background, that *gpt*-SV40 sequences are detectable on this blot also (the probe also hybridises to the pSV_1ov backbone in the control tracks, by virtue of complementary SV40 sequences).

These results were the first indication that the cointegration frequency of the non-selected gene with the selected gene, following simple cotransfections, is unsatisfactory. This is discussed in greater detail in section 4.2.3.



AUTORAD. A
(gpt-SV40 probe)

AUTORAD. B
(ovalbumin probe)



AUTORAD. C
(gpt-SV40 probe)

some of the transfectants contain a single *gpt*-sequence. If there are multiple integrated copies of the plasmids arranged in a concatameric unit (tandem head-to-tail repeat), a genomic digest with a restriction endonuclease which cuts once within the plasmid sequence should generate a hybridising fragment which is the same size as the full-length plasmid. In fact, in autoradiogram C, only the pTK-E(b')gpt cotransfectant, clone C9, which appears to contain multiple integrated-copies of *gpt* sequence (see autoradiogram A) gives a fragment around the full size of 7.3Kb (the pML-Egpt cotransfectants also give bands corresponding to the molecular weight of the linearised fragment (6.2Kb) but these are very faint and only just visible on the autoradiogram). Hence, it might be concluded that some of the other cotransfectants, whose genomic DNA gives large-size hybridising fagments (in autoradiogram C), contain single copies of integrated plasmid. However, it is equally possible that the large fragments represent only partially-digested concatameric sequences. In addition, this interpretation is not consistent with results of autoradiogram A, where the copy number of integrated *gpt* sequences appears variable. In fact, the estimated copy number of the integrated sequences from autoradiogram A seems different to the copy numbers estimated from autoradiogram C. This might be due to unequal washing of the probe on various sections of the blot which produced autoradiogram C, as suggested by the variable background.

In summary, the blots of figure 4.2. are ambiguous in terms of revealing details of the integration pattern of the DNA. At this early stage of the investigation, however, the aim was simply to identify the cointegration of the ovalbumin cDNA with the *gpt* sequence. This event was not detectable in the six clones screened. Although by no means conclusive, this result led to the suggestion that cotransfection (introduction of both plasmids into the host cell) by spheroplast fusion is not efficient and, therefore, there is little opportunity for cointegration of the two plasmids to occur.

One suggested explanation for spheroplast-fusion cotransfections being inefficient is that they require fusion of a cell with at least two spheroplasts, and this three-way fusion might be an infrequent event. There were two means of circumventing this potential problem. An alternative transfection method could be used, or the spheroplast-fusion approach could be adapted such that spheroplasts carried both plasmids. The latter option

requires manipulation of the plasmids, such that they contain different antibiotic-resistance genes to allow selection for their dual existence in one bacterial cell. Also, the plasmids require different copy-number-control sequences, if the excess of the plasmid carrying the non-selectable gene is to be maintained. However, phenomena such as plasmid incompatibility, which are not well understood and are difficult to forsee, might present problems with such an approach. In addition, the observation that following a spheroplast-fusion transfection, there is little residual viscosity in the medium, suggests that many of the spheroplasts are taken up into the cells. Otherwise, they would lyse and the released DNA would be expected to make the medium viscous. It was, therefore, considered more straightforward to choose an alternative transfection-procedure for simple cotransfections, as described below.

4.2.4. Simple cotransfection using electroporation.

A cotransfection between pMMTV-Eova and pMMTV-Egpt (see figures 4.1. and 3.6. respectively) was performed, using electroporation (see section 2.3.2.) as an alternative transfection method to spheroplast fusion. The total concentration of plasmid DNA used was 0.12 mg/ml, with a copy-number ratio of pMMTV-Eova : pMMTV-Egpt (10:1). Both plasmids were linearised at a *Nde* I site, which is situated in the pBR-backbone sequence, before adding to the cells. The stable *gpt*$^+$-transfection frequency obtained was 1.2 x 10^{-4} transfectants per cell, which is higher than might be expected for such a cotransfections using spheroplast fusion (see table 3.1.). This may suggest that at least as much DNA was introduced into the cells during the electroporation as in spheroplast-fusion transfections. Genomic DNA was prepared from stable *gpt*$^+$-transfectant clones and analysed by Southern blotting (see figure 4.3.). As expected, in *Hind* III / *Bam* HI genomic digests, the 1.9Kb *gpt*-SV40 fragment present in pMMTV-Egpt was detectable in all genomes using the *gpt*-SV40 probe. When the filter was stripped and rehybridised to the ovalbumin probe, the ovalbumin cDNA positive-control fragment gave a more intense signal than that from the same amount of the *gpt*-SV40 control sequence, yet the ovalbumin sequence (a 1.6Kb *Hind* III - *Bam* HI ovalbumin-cDNA - SV40 fragment was exected) was not detectable in the genomic tracks. As explained in the figure legend, it

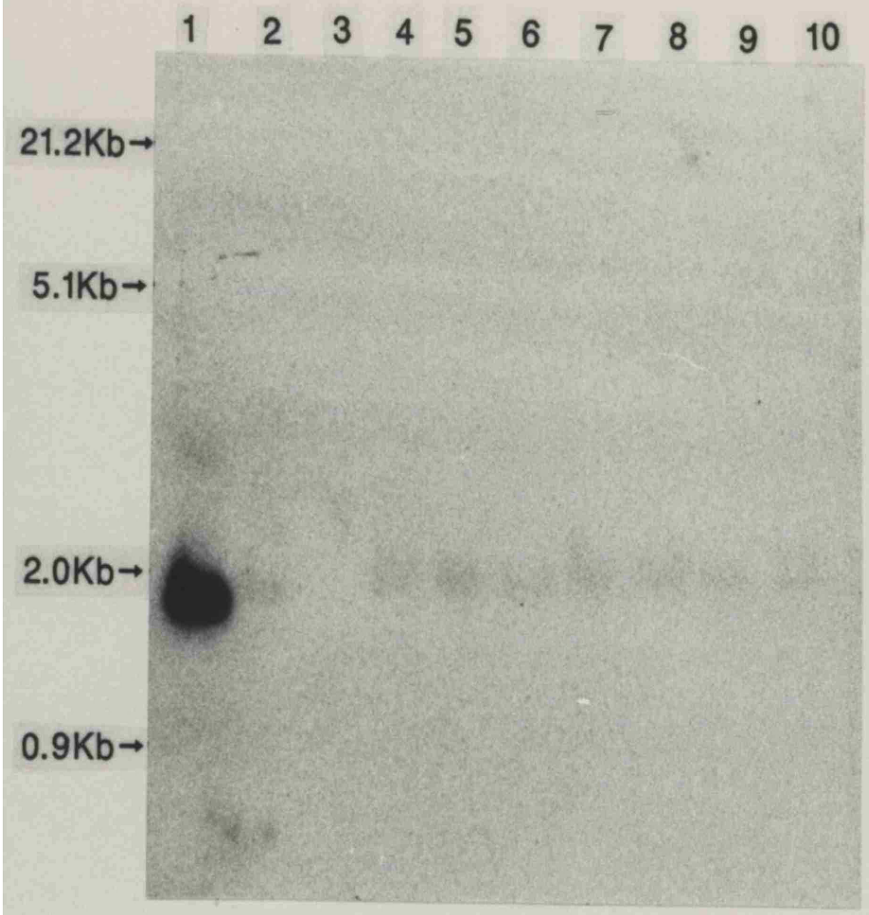
***Figure 4.3. Southern blot to analyse for cointegration of the ovalbumin-cDNA sequence with the gpt* sequence following simple cotransfection by electroporation.**

pMMTV-Eova and pMMTV-Egpt were cotransfected into J558L cells by electroporation (see section 4.2.4.). Genomic DNA was isolated from fourteen, monoclonal, stable *gpt*$^{+}$-transfectants. These DNA preparations were then digested with *Hind* III and *Bam* HI, run out on a 1% agarose gel and Southern blotted (see section 2.11.3.). The blot was first hybridised to the *gpt*-SV40 probe (autoradiogram A), then stripped and hybridised to the ovalbumin probe (autoradiogram B). These probes are described in section 2.10.1.

The tracks of the blot shown contained the following samples: tracks 1 and 2: *Hind* III / *Bam* HI digests, of 10μg untransfected-J558L genomic DNA, and of pSV_2gpt, containing 150pg and 1.5pg respectively of the 1.9Kb *gpt*-SV40 fragment; track 3: *Hind* III / *Bam* HI digests, of 10μg untransfected-J558L genomic DNA, and of a plasmid, pSV_2ov, containing 1.5pg of the 1.4Kb ovalbumin-cDNA fragment ; tracks 4-10: 10μg of *Hind* III / *Bam* HI digested genomic-DNA from seven of the cotransfectant clones (which are typical of the analysis as a whole). Tracks 1-3 serve as positive or negative controls, depending on the probe being used.

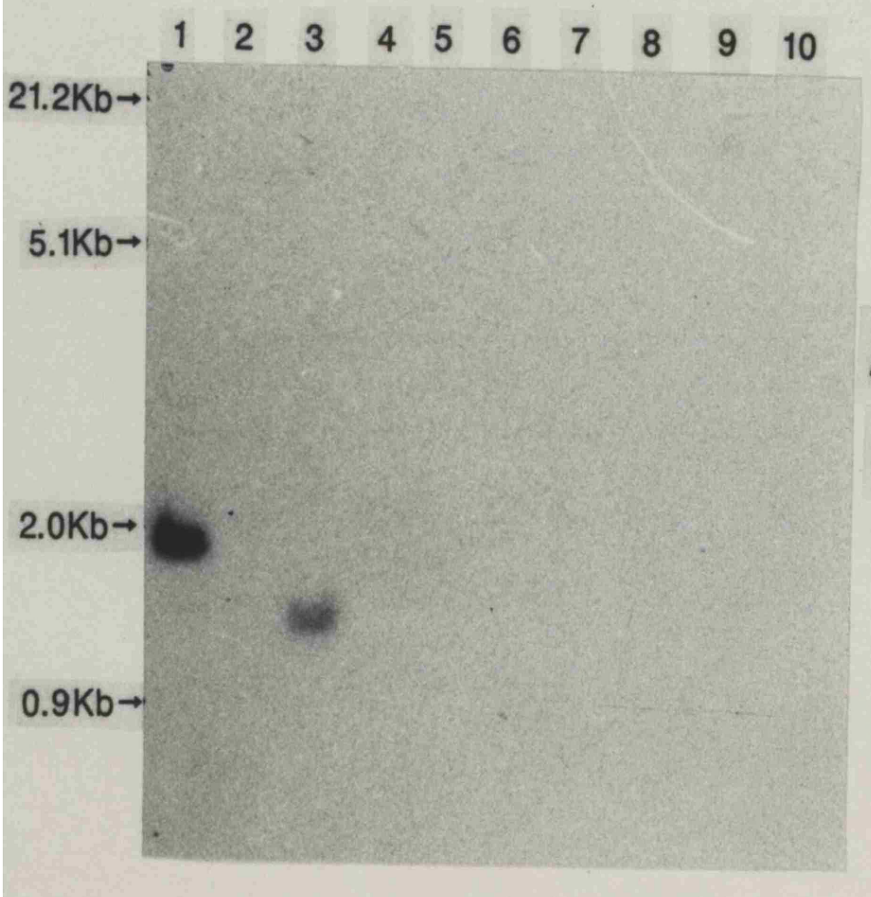
Autoradiogram A shows that in all transfectants, the 1.9Kb *Hind* III - *Bam* HI *gpt*-SV40 sequence is detectable in the genomic DNA. The copy number of the integrated sequences appears to be low. Assuming 6 x 10^{9} bp per diploid genome, if there is one copy per genome of the 1.9Kb *gpt*-SV40 fragment, there is 3.2pg of this sequence in 10μg of genomic DNA; the genomic tracks give similar or only slightly stronger signals than that from 1.5pg of the same 1.9Kb fragment in the control track, track 2, suggesting that the *gpt*-SV40 sequence is integrated in most transfectants as a single copy and, possibly, in some transfectants with two copies.

After stripping the blot, and then hybridising to the ovalbumin probe, no integrated ovalbumin sequences are detectable (autoradiogram B). This is despite the strong signal from 1.5pg of the 1.4Kb ovalbumin-cDNA fragment from pSV_2ov, in control track 3 (there is also residual hybridisation of the *gpt*-SV40 probe in track 1). This result suggests that the poor cointegration-frequency of the non-selected gene with the selected gene, following simple cotransfection, is not limited only to spheroplast fusion, but also to electroporation.



AUTORADIOGRAM A

(gpt-SV40 probe)



AUTORADIOGRAM B

(ovalbumin probe)

seems that in all or most of the transfectants, the *gpt* sequence was integrated as a single copy. This is consistent with the observation of Potter *et al.* (1984) who reported that the gene copy-number of transfected sequences following electroporation is generally low. As the detection level is down to single copy, this shows that no cointegration of the ovalbumin-cDNA sequence occurred with the *gpt* sequence in the fourteen stable *gpt*$^+$-transfectants screened.

4.2.5. Discussion of simple cotransfection.

I was unable to detect cointegration of the non-selectable ovalbumin gene with the *gpt* gene, in a score of *gpt*$^+$ transfectants, which were produced using either spheroplast fusion or electroporation. Hence, the approach of simple cotransfection was not found to be a satisfactory method for the routine production of J558L transfectants carrying a stably-integrated non-selected gene. Expanding and analysing individual colonies is a very time consuming and labour intensive process. Therefore, it is desirable that a high proportion (ideally approaching 100%) of the transfectants cointegrate the non-selected gene with the selected gene, so as to reduce the number of clones that have to be screened for expression of the non-selected-gene product. One potential strategy for overcoming the low cointegration frequency was to increase the ratio of the non-selectable gene to the *gpt* gene in the cotransfections. However, the results described above suggest that the frequency of cointegration is so low, that the ratio might have to be increased to a point where the number of *gpt* genes introduced in transfections is so small (particularly in spheroplast fusions), that stable *gpt*$^+$-transfectants are no longer generated. Gopal (1985) emphasised that difficulty in obtaining satisfactory cointegration-frequencies following simple cotransfection is mainly confined to non-adherent cells, of which myeloma cells are a prime example. He reported a novel transfection protocol, which involves the attachment of cells to a concanavalin-A - coated cell-culture dish, followed by treatment of cells with DEAE-dextran to allow adsorption of plasmid DNA to the attached cells, and finally treatment with a solution of polyethylene glycol to facilitate the uptake of DNA by the cells. In unpublished results, he reported cointegration frequencies as high as 75%, perhaps this procedure could also

be successfully used in J558L. However, I used the strategy described in section 4.3. to increase the cointegration frequency.

4.3. Cointegration via a two-gene plasmid.

Studies on transfected cells usually indicate that all or most of the sequences of a passively-replicated expression plasmid (see section 1.5.1.1.) become incorporated into the host genome during an integration event. Therefore, it is predictable that if a selectable gene and a non-selectable gene reside on the same plasmid, their cointegration frequency will be much higher than with simple cotransfections.

pSV-Vμ1 (figure 4.4.) is a plasmid which contains two genes on the same plasmid. It is a pSV_2gpt derivative, in which the *Bam* HI - *Eco* RI fragment, which contains SV40 late-region sequences, is substituted with a mouse IgH-gene (V_H-Cμ). Although this gene was constructed by recombinant methodology (see Neuberger, 1983), its structure is typical of a complete, rearranged IgH(μ)-gene. pSV-Vμ1 was shown to give rise to μ-chain expressing transfectants in several B-lymphoid cell lines (Neuberger, 1983). Providing the μ gene can be integrated, this is a predictable result, because the μ gene contains expression signals which are compatible with such cell-types. Likewise, if integration of this μ gene into the genome of J558L could be obtained following transfection with pSV-Vμ1, it was expected that μ chain would be expressed.

J558L cells were transfected with pSV-Vμ1 by spheroplast fusion (see section 2.3.1.) using the *E. coli* strain, DH5, as the bacterial host. Initially, μ-chain secretion was analysed for from stable gpt^+-transfectant polyclonal populations (≈ 30 clones per population). The proteins of these cell populations were biosynthetically-labelled with ^{35}S-methionine (see section 2.5.1.), and cell supernatants were then analysed for expression and secretion of the transfected - μ-gene product (μ-chain), by immune-precipitation (see section 2.5.2.) from the cell supernatant, using commercially-supplied rabbit anti - mouse-IgM antibody. Immune-precipitates were run out on SDS-PAGE gels (see section 2.5.3.). Autoradiograms (not shown) of the gels detected μ-chain secretion from 19 of 20 populations screened. This result seemed to confirm that cointegration of a non-selected gene with a selected gene occurs at a much higher frequency when

Figure 4.4. Partial restriction map of pSV-Vμ1.

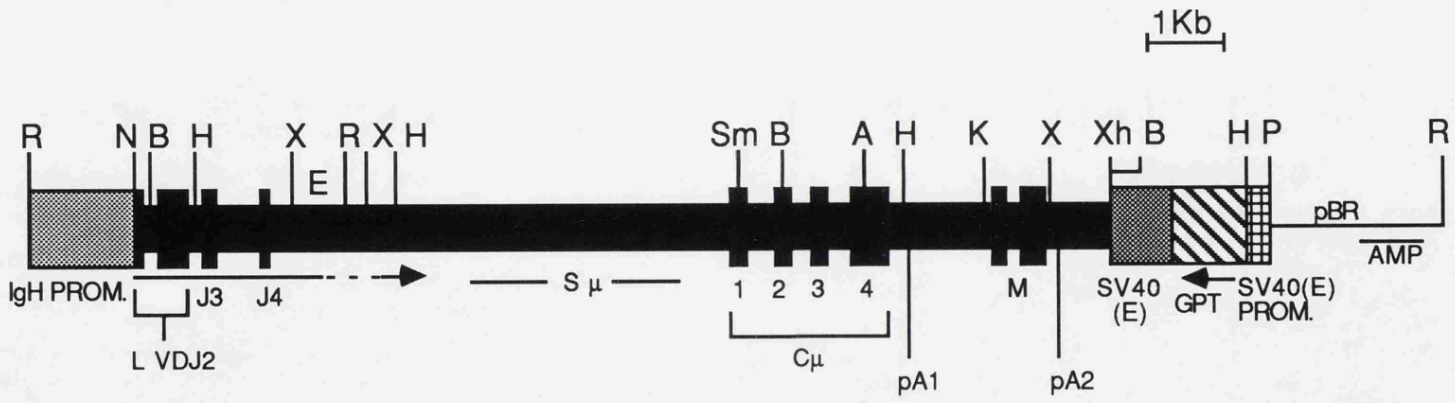
A partial restriction map of pSV-Vμ1, showing the sites relevant to the manipulations made in this study, is shown. The plasmid contains two converging transcription units, one drives *gpt*-transcription and the second is a complete rearranged IgH(μ)-gene.

The IgH promoter is represented by a box filled with dots. The black shaded areas represent potentially transcribed μ-gene sequences; the thinner boxes represent introns, including the major intron (exons J_3 and J_4 are also spliced out as part of the major intron), which contains the 1Kb *Xba* I enhancer fragment and the switch region (Sμ); the thicker boxes represent the exons, namely, the leader sequence (L), the fused variable, diversity and joining regions (VDJ_2), the constant region exons (Cμ; 1, 2, 3 and 4) and membrane region exons (M). The SV40 early-promoter fragment (SV40(E) PROM.) [also containing the associated enhancer and late promoter] is represented by a box filled with squares, the *gpt* sequence is represented by a box filled with thick diagonal lines, the SV40 sequence downstream of *gpt* (containing the small t-antigen intron and the early polyadenylation-signal (SV40(E))) is represented by a darkly shaded box and the pBR backbone, which contains the ampicillin-resistance gene (amp), is represented by a thin line.

The positions of the two polyadenylation signals in the 3'-region of the μ-gene are indicated. Polyadenylation at pA_1 leads to production of μ_s mRNA, and polyadenylation at pA_2 leads to production of μ_m mRNA (see section 1.12.3.). In J558L, a myeloma cell-line, pA1 is expected to be preferentially used.

The plasmid is shown as linearised at the *Eco* RI site. Restriction sites are abbreviated as follows: *Hind* III: H, *Nco* I: N, *Bam* HI: B, *Xba* I: X, *Eco* RI: R, *Sma* I: Sm, *Apa* I: A, *Kpn* I: K, *Xho* I: Xh, *Pvu* II: P.

pSV-Vμ1


1Kb
R
N B H
X R X H
E
Sm B
A H
K
X
Xh B
H P
R
pBR
AMP
IgH PROM.
J3
J4
L VDJ2
S μ
1 2 3 4
Cμ
pA1
M
pA2
SV40
(E)
GPT
SV40(E)
PROM.

they are physically linked.

Monoclonal cell populations were derived from two of the polyclonal populations secreting detectable levels of μ chain, by cloning (see section 2.2.3.). Their relative levels of μ-chain secretion were estimated, by performing biosynthetic-labelling and immune-precipitation, and then determining the relative amounts of μ chain in immune-precipitates by scintillation counting (see section 2.5.4.). Figure 4.5. illustrates that 25-50% of the monoclonal transfectants gave counts significantly above background, and that a particularly high-level expressing clone, clone D4, was identified. As is usually the case in a population of transfected cells, the secretion level of the transfected-gene product (μ chain) was highly variable. As discussed in section 1.8., this is probably due to a combination of the gene copy-number and the influence of surrounding chromosomal sequences. This result illustrates the importance of screening transfectants individually in order to obtain high-level expression of a transfected-gene product.

4.4. Stable expression from a transfected expression-plasmid in J558L.

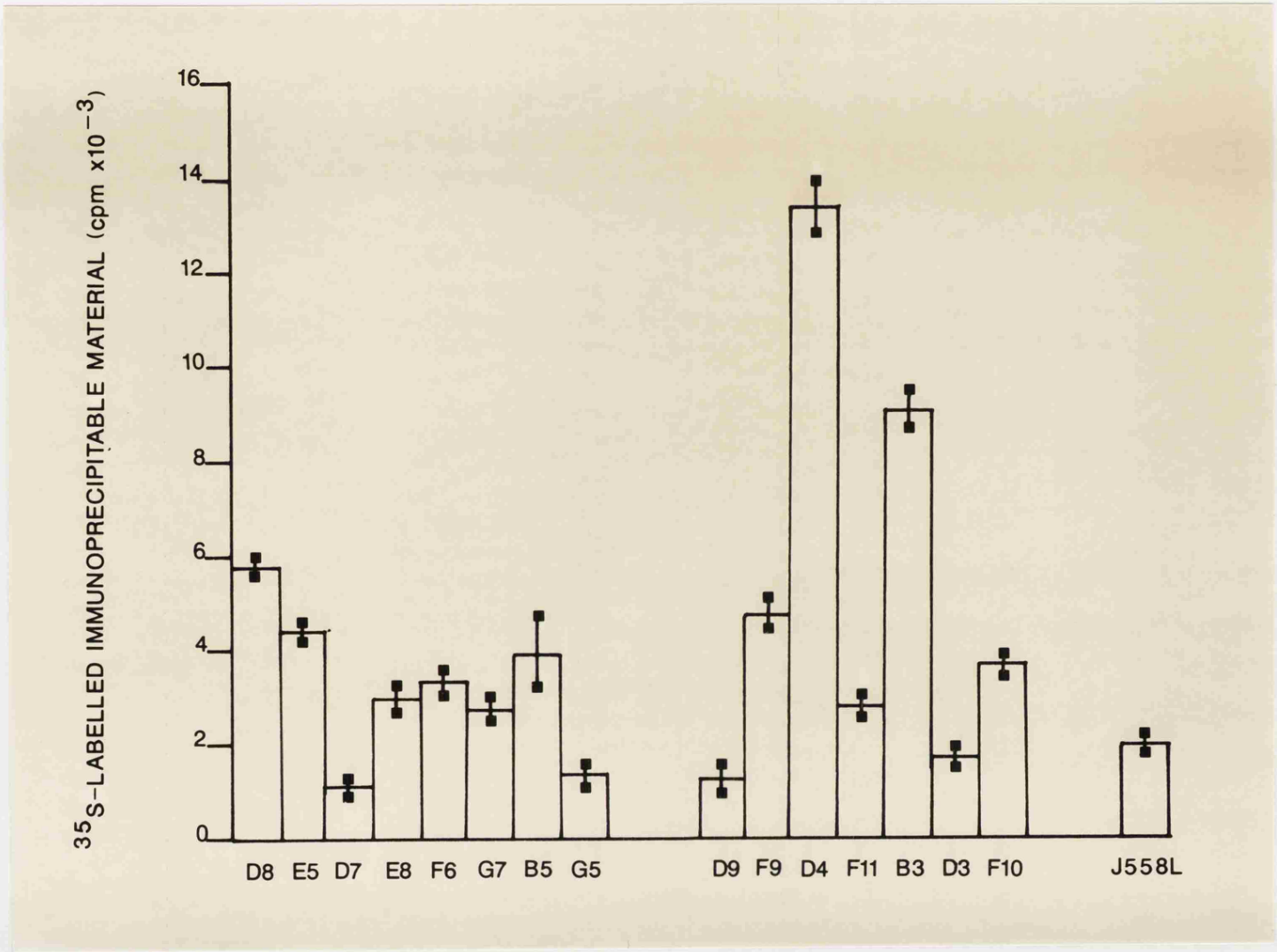
It was decided to investigate if expression from transfected sequences could be maintained in the absence of selection. Some vectors are notoriously unstable such as those based on the BPV replicon (see section 1.5.1.2.). Integrated sequences may also show instability. For example, Gebara *et al.* (1987) showed that when stable *gpt*$^{+}$-transfectants of mouse fibroblasts were grown in non-selective medium, they reverted (lost expression) at a low frequency (≈ 10^{-4} per cell division). Numerous mechanisms, both genetic and epigenetic, were identified as being responsible for the inactivation. The instability of integrated sequences seems to be both sequence and host-cell dependent. For example (as discussed in section 1.5.2.1.), some integrated Retroviral-sequences are particularly unstable.

The highest-level μ-chain expressing pSV-Vμ1-transfectant of J558L, clone D4 (see figure 4.5.), was grown in non-selective medium for four weeks and then subcloned (see section 2.2.3.). Six subclones were expanded and after growth for a further total of 10 weeks in non-selective medium, the relative levels of μ-chain secretion were compared to see if

Figure 4.5. Screening of pSV-Vμ1 transfectants of J558L for μ-chain secretion, and isolation of a high-level secreting transfectant.

Monoclonal cell-populations were derived from two pSV-Vμ1-transfectant polyclonal populations of J558L (each containing about 30 pooled colonies), by subcloning. Proteins of these cell lines were biosynthetically-labelled with ^{35}S-methionine, and then the relative levels of μ chain (the product encoded by the μ gene on the transfected plasmid, pSV-Vμ1) secreted were determined, by immune-precipitation from equal volumes of cell supernatants with rabbit anti-mouse IgM antibody followed by scintillation counting. Immune-precipitations were performed in duplicate on each supernatant.

The histogram shows that the level of μ-chain secretion was highly variable between clones. This is an observation also made in later analyses of lysozyme secretion from transfected cells (e.g. see figure 4.25.). Four of the fifteen clones analysed gave counts slightly lower than the background count of an 'immune-precipitate' from an untransfected-J558L supernatant (negative control). This may be because not all cells were in optimal condition during this early stage of screening. Altogether, it appears that at least half the clones were secreting insignificant amounts of μ chain. However, the analysis was sufficient to allow isolation of a high-level expressing clone, clone D4, which was further characterised, as is described in sections 4.4. - 4.6.


^{35}S-LABELLED IMMUNOPRECIPITABLE MATERIAL (cpm x10^{-3})
16
14
12
10
8
6
4
2
0
D8
E5
D7
E8
F6
G7
B5
G5
D9
F9
D4
F11
B3
D3
F10
J558L

the levels of secretion showed any variation. Figure 4.6. shows that the secretion levels from the subclones were quite similar, except from one subclone (B1), in which the secretion level was much lower than in all other clones. Subclone A1 gave the highest expression, although the difference was only about two-fold compared with most of the other subclones. Such minor differences in expression levels are often seen following cloning. This might be due to the naturally occurring variability in expression of a protein from cells of a cell lineage, which only becomes apparent following cloning, when all cells express similar levels of any one protein. Analysis of clone-D4 cells, grown in selective medium to preserve the level of expression, was not included in the experiment. Therefore, it is not possible to exclude the possibility that there was a universal change in the μ-chain expression level. However, the stability of expression levels from transfected genes in J558L transfectants seemed to be supported by studies with a lysozyme-expressing transfectant, clone D8 (see section 4.7.6.). This was grown in non-selective medium for several months with no detectable change in the lysozyme secretion-level, relative to the level of IgG secretion from the myeloma cell-line, X63.Ag8. Clone D8 was also reintroduced into selective medium without any visible effects on growth, suggesting that expression from the *gpt* gene was also maintained. Ideally, in order to demonstrate the stability of expression from the *gpt* gene, the cells would be grown for a longer period in non-selective medium, and then their cloning efficiencies would be compared in selective (XHMPA supplemented) and non-selective media; if the cloning efficiencies were similar, this would suggest that the cells still express the *gpt*-gene product, XGPRT, which allows them to survive selection.

At the small-scale level of laboratory cell-culture work, there is usually no need to remove selection. However, in large-scale culture, addition of the selecting agents, particularly for some other selection systems such as those based on the DHFR or adenosine-deaminase genes (see sections 1.7. and 1.8.), may add considerably to costs. Therefore, it is desirable that the transfectants are sufficiently stable to allow removal of the selecting agents following the initial selection. As already alluded to, in the above described experiment, the stability of the cells in non-selective medium would ideally have been assessed over a longer time-period. Even so, the results suggest that gpt^+ transfectants of J558L are

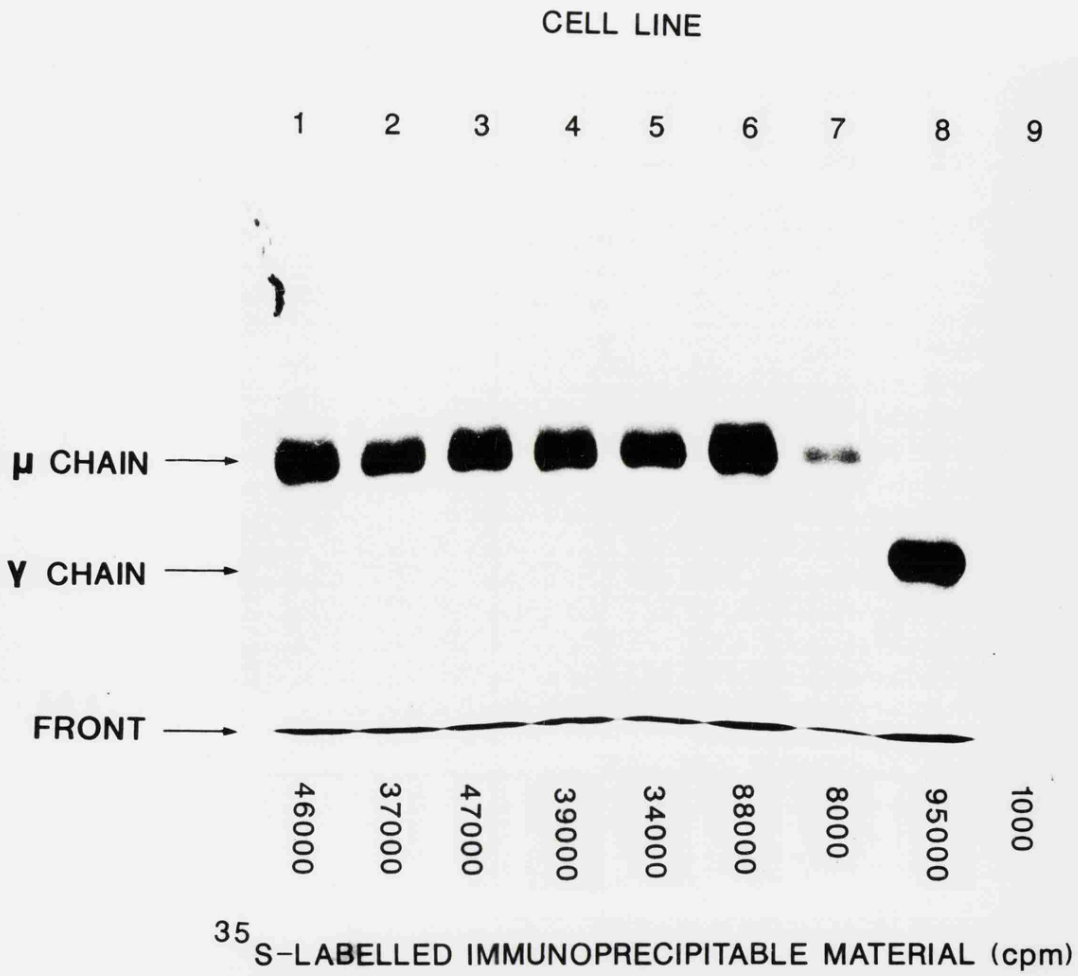
Figure 4.6. Stability of μ-chain expression in a pSV-Vμ1 transfectant of J558L.

Clone D4 was shown to express the highest level of μ-chain amongst the fifteen monoclonal pSV-Vμ1-transfectants of J558L screened (see section 4.3. and figure 4.5.). This cell line was grown in non-selective medium for four weeks and then subcloned. Seven derivative clones were expanded. After a further total of ten weeks growth in non-selective medium, their relative levels of μ-chain secretion were compared; cell protein was biosynthetically-labelled with ^{35}S-methionine, and μ chain was immune-precipitated from cell supernatants. All quantitations were made on the same volume of cell supernatant following labelling under identical conditions. The immune-precipitates were analysed by SDS-PAGE (8%) and autoradiography.

The autoradiogram shown demonstrates the apparent specificity of the μ chain immune-precipitations when the level of expression is high. Hence, μ chain (actually IgM; see figure 4.7.) could also be quantified by scintillation counting of immune-precipitates, in the knowledge that the quantitation is not significantly affected by background. The counts from such analyses are indicated below each track, and these correlate well with the relative intensities of the μ-chain bands.

The tracks contain the following samples: tracks 1-7: μ chain from supernatants of the following subclones of clone D4, immune precipitated with rabbit anti - mouse-IgM antibody:- track 1: clone A6; track 2: clone C6; track3: clone E5; track 4: clone C3; track 5: clone A2; track 6: clone A1; track 7: clone B1; track 8: γ chain from the supernatant of myeloma cell-line, X63.Ag8, immune-precipitated with rabbit anti - mouse-IgG antibody; track 9: rabbit anti - mouse-IgM and rabbit anti - mouse-IgG antibodies (1:1(v/v)) 'immune-precipitate' from untransfected-J558L supernatant (negative control).

The analysis suggested (as discussed in section 4.4.) that in the absence of selection, the transfectants are reasonably stable, because only one subclone of clone D4, (B1), secreted a vastly different level of μ chain.


CELL LINE
1
2
3
4
5
6
7
8
9
μ CHAIN
γ CHAIN
FRONT
46000
37000
47000
39000
34000
88000
8000
95000
1000
35S-LABELLED IMMUNOPRECIPITABLE MATERIAL (cpm)

sufficiently stable to be grown in non-selective medium over a three-month period, which is typical of the time scale required to build up large-scale 'continuous' cultures and harvest their products. The indication that subclone B1 underwent a reduction in μ-chain expression levels, however, emphasises the importance of maintaining starter-cultures in selective medium.

4.5. The protein encoded by a transfected gene associates with an endogenous protein.

The SDS-PAGE gel in figure 4.6. contained only 8% acrylamide and, therefore, retained only the IgH chains. But an 11% gel (see figure 4.7.) illustrates that μ chain, produced in a pSV-Vμ1 transfectant of J558L (clone D4:A1), is found in the supernatant, complexed with what appears to be λ light-chain, presumably in the form of a complete Ig molecule (μ_2,λ_2). Although J558L is a heavy-chain-loss variant, it secretes the endogenous λ light-chain (Oi *et al.*, 1983; also see figure 4.16.). However, this λ chain was not recognised by the rabbit anti - mouse-IgM antibody. In contrast, a band corresponding to the size of λ light-chain was immune-precipitated from the supernatant of the pSV-Vμ1 transfectant by rabbit anti - mouse-IgM antibody. The result from this experiment suggests the antibody only recognised the μ chain expressed from the transfected μ-gene and, therefore, the λ chain was coprecipitated with the μ chain. Unless assembly of this Ig complex occurs post-secretion, then assembly between the endogenous(λ)- and exogenous(μ)-gene products must occur within the secretory pathway.

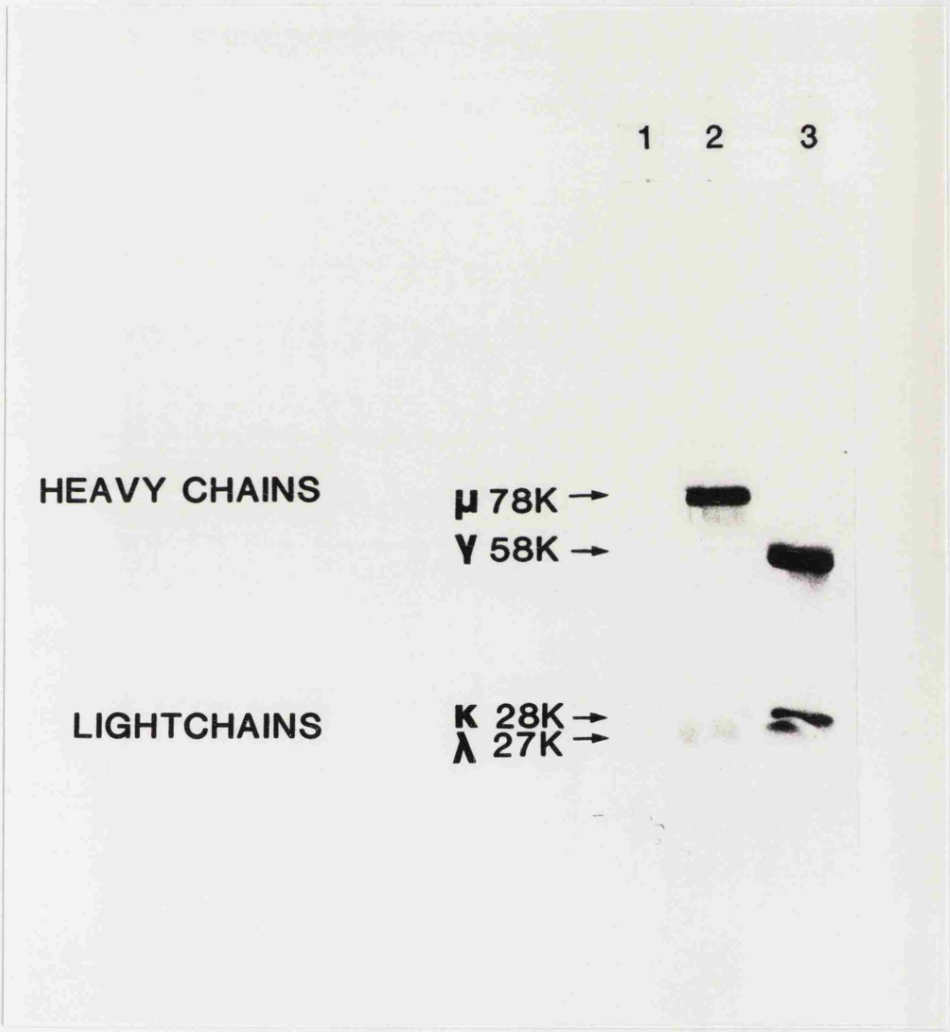
The above result is one which might be expected because, in some cases, complexing of IgH and IgL chains seems to be essential as a pre-requirement for their secretion. For example, heavy-chain-loss variants of the S107 myeloma cell-line do not secrete the endogenous light chain (Oi *et al.*, 1983). Also, IgH (μ) chain may accumulate in the ER unless able to complex with an IgL chain (Mains *et al.*, 1983). There is evidence that unassembled or unfolded chains are retained in the ER until assembly and folding is completed. This may be partially explained by the binding of IgH chains to a 77KD protein called BiP (*bi*nding *p*rotein) (Haas and Wabl, 1983; Bole *et al.*, 1986), which is thought to play a role in the oligomerisation

Figure 4.7. SDS-PAGE analysis of secreted Ig from a pSV-Vμ1 transfectant of J558L.

Figure 4.6. demonstrated the secretion of μ chain from subclones of the pSV-Vμ1 transfectant, clone D4, on a SDS-PAGE gel. However, the gel contained only 8% acrylamide and, therefore, only IgH chains were retained on the gel. Hence, the analysis was repeated, except this time using an 11% gel, which also retains IgL chains. The photograph shows the autoradiogram from such a gel.

The tracks contain the following samples: track 1: rabbit anti - mouse-IgM antibody 'immune-precipitate' from untransfected-J558L cell supernatant (negative control); track 2: rabbit anti - mouse-IgM antibody immune-precipitate from the cell supernatant of a pSV-Vμ1 transfectant of J558L (A1 subclone of clone D4); track 3: rabbit anti - mouse-IgG antibody immune-precipitate from X63.Ag8 cell-supernatant. The approximate sizes of the immune-precipitated proteins, as determined from a molecular-weight protein marker migration-curve, are indicated.

As discussed in section 4.5., this autoradiogram suggests that the μ chain encoded by the transfected μ-gene, is secreted as a complex, with what appears to be endogenous, λ light-chain. The λ chain gives a weak signal because of its low methionine-content (1 residue per chain).


1 2 3
HEAVY CHAINS
μ 78K
γ 58K
LIGHTCHAINS
κ 28K
λ 27K

and folding process.

The requirement for association of Ig chains prior to secretion has been exploited for the production of chimeric antibodies with novel effector-functions (reviewed by Neuberger, 1985; Williams, 1988). This involves DNA manipulation of Ig genes prior to their introduction into lymphoid cells, followed by the assembly of the transfected-Ig-gene product with other Ig chains encoded by endogenous or exogenous genes, .

4.6. High-level expression of IgH chain from a transfected IgH-gene.

Figure 4.7. showed immune precipitates of Igs secreted from the myeloma cell-lines, X63.Ag8 and clone D4(A1) (a pSV-Vμ1 transfectant of J558L). It appeared that the levels of Ig secretion from the two cell lines were similar. Hence, biosynthetic-radiolabelling experiments were performed under the same conditions, in triplicate, on both cell lines, after carefully maintaining the cells in good condition for several days. The secreted Igs were immune-precipitated from cell supernatants, and their relative levels were estimated by scintillation counting (see figure 4.8.). Taking the methionine content of the Ig chains into account, it was estimated that the relative molar-expression of Ig molecules from the J558L transfectant, clone D4(A1), was 9/10 of that from the control myeloma-cell, X63.Ag8. The calculation is explained in table 4.1.

The level of IgG secretion from X63.Ag8 is considered to be fairly typical of the Ig secretion-level from myeloma cells in general. Hence, the above result suggests that transfectant clones can be isolated in which expression from a transfected gene is as efficient as from an endogenous Ig-gene. If so, it shows that negative influences on expression, such as the site of chromosomal integration, can be overcome using the expression system described, providing enough transfectant clones are screened. However, it would be wrong to infer, from this evidence, that the protein encoded by any gene, whose transcription is regulated by an IgH enhancer and a promoter and which is present on a plasmid also containing a *gpt* gene, can be expressed in stable *gpt*$^{+}$-transfectants at such high levels; expression from a transfected Ig-gene in a myeloma cell can hardly be recognised as a truly heterologous system, because specific post-transcriptional control mechanisms seem to operate to ensure high-level Ig

Figure 4.8. Comparative Ig-secretion levels from X63.Ag8 and the pSV-Vμ1 transfectant of J558L, clone D4(A1).

The proteins of optimally-growing cells were biosynthetically-labelled with ^{35}S-methionine. The amount of Ig secreted into the medium was determined, by immune-precipitation from equal volumes of cell supernatants followed by quantitation of immune-precipitates by scintillation counting. Labelling experiments were performed in triplicate for each cell-line. IgM secreted from the pSV-Vμ1 transfectant of J558L, clone D4(A1), was immune-precipitated with rabbit anti - mouse-IgM antibody, IgG secreted from X63.Ag8 was immune-precipitated with rabbit anti - mouse-IgG antibody and an 'immune-precipitation' was performed on untransfected-J558L supernatant with a mixture (1:1(v/v)) of rabbit anti - mouse-IgG and rabbit anti - mouse-IgM antibodies (negative control).

The histogram illustrates the average number of immune-precipitable counts obtained from each cell-line. The small filled squares indicate the range in the determinations for each cell-line. The results of this quantitation were used to estimate the relative molar-levels of Ig secretion, as explained in table 4.1.

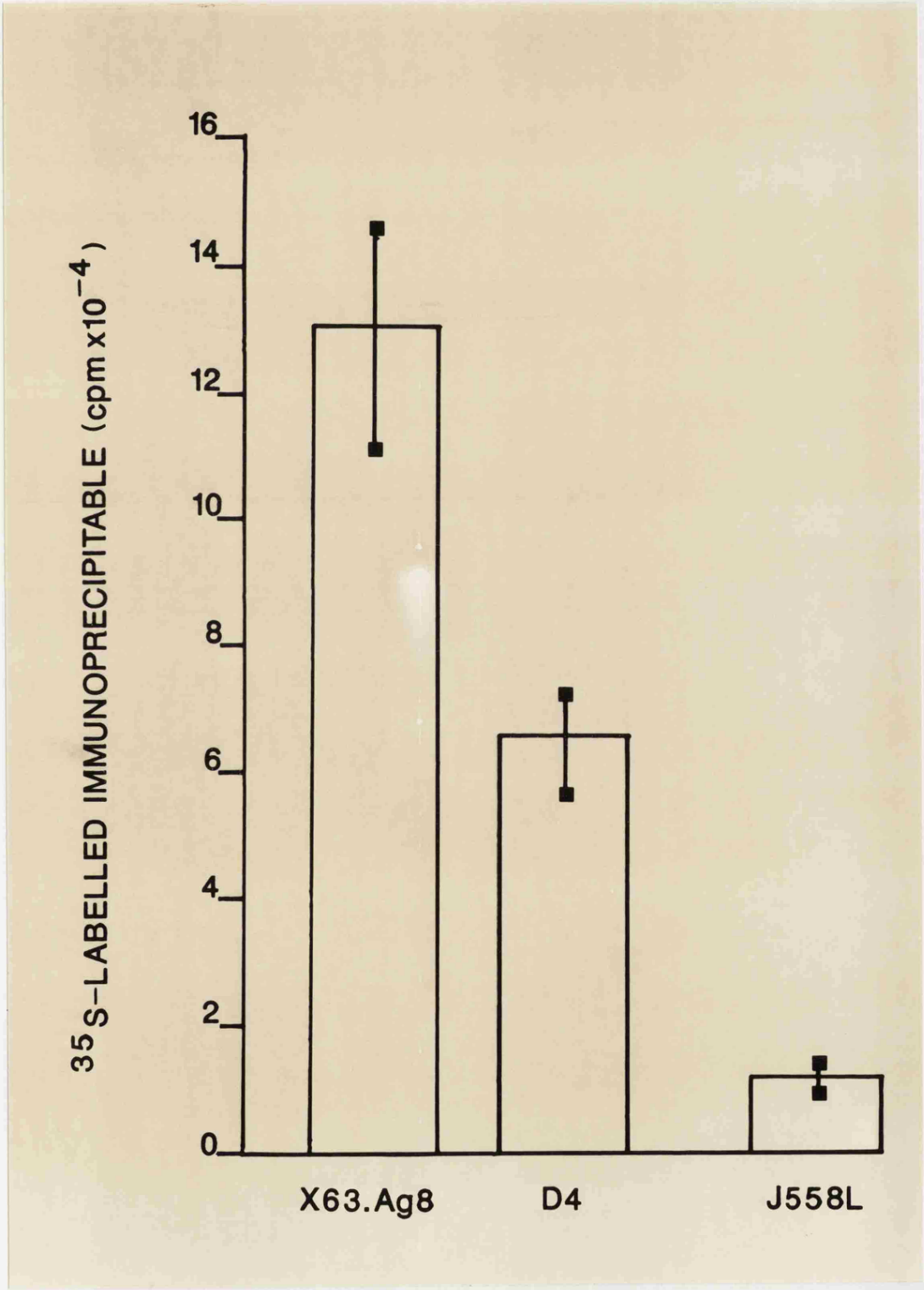

^{35}S-LABELLED IMMUNOPRECIPITABLE (cpm x10^{-4})
16
14
12
10
8
6
4
2
0
X63.Ag8
D4
J558L

Table 4.1.. Estimation of relative molar Ig-secretion from X63.Ag8 and the pSV-Vμ1 transfectant of J558L, clone D4(A1).

The average determinations of immune-precipitable material from cell supernatants of X63.Ag8 and clone D4(A1) (see figure 4.8.) were used to estimate the relative molar-levels of Ig secretion from the two cell lines, as shown in the table below.

The number of counts from scintillation analysis of immune-precipitates reflects not only the amount of Ig secreted, but also the methionine content of the Ig chains. The methionine content of the MOPC 21 (Köhler and Milstein, 1975) was obtained from reports by Svasti and Milstein (1972) (for κ chain) and by Adetugbo (1978) (for γ_1 IgH chain), and the methionine content of the μ and λ chains was obtained from Kabat *et al.* (1983). Assuming that both immune-precipitations give quantitative recovery (this is shown for the recovery of $IgG_1(\kappa)$ from a X63.Ag8 cell-supernatant, see figure 4.16.), and that all Ig secreted from the pSV-Vμ1 transfectant, clone D4(A1), is in the form of the tetramer IgM, μ_2, λ_2, then the molar Ig-secretion levels from the two cell lines are approximately equivalent. This demonstrates that the product of a transfected gene can be expressed as efficiently as the highest-expressed endogenous protein of a myeloma cell.

Cell line	Ig secreted	Immunoprecipitable material (cpm)	Met. res./ Ig molecule	Relative molar expression
X63.Ag8	$(\gamma_1)_2$, κ_2 (MOPC 21)	129467	26 (γ=9, κ=4)	1.0
D4(A1)	μ_2, λ_2	64241	14 (μ=6, λ=1)	0.9

expression in myeloma cells (see section 1.12.). Therefore, it still remained to be demonstrated that a heterologous-gene product, which is not normally produced in myeloma cells, can be expressed at similar levels to that of Ig molecules.

4.7. Divergent two-gene expression plasmids.

The results obtained from transfections with pSV-Vμ1 suggested that the cointegration frequency of a non-selected gene with a selected gene is high when both genes are present on the same plasmid. Therefore, expression plasmids were constructed which contain a *gpt* transcription-unit (almost identical to that found in the pPROM-Egpt plasmids; see section 3.3.1.), a second transcription-unit for a heterologous gene of interest and the IgH-enhancer sandwiched between the two transcription-units. The two transcription-units have opposite orientations and are, therefore, expected to drive transcription in opposite directions away from the enhancer. For this reason, they are called divergent expression-plasmids.

4.7.1. Plasmid construction.

The flanking sequences for a second transcriptional-unit were built into the pPROM-Egpt construct, pΔTK-Egpt (see section 3.3.1.), to generate a plasmid called pΔTK-Egpt-CMV (see figure 4.9.). pΔTK-Egpt-CMV is a convenient plasmid for inserting any non-selectable cDNA sequence, from which expression in myeloma cells is required. It contains an IgH enhancer, which separates the *gpt* transcription-unit from the Human Cytomegalovirus immediate-early region (CMV-IE) promoter and poly(A) sequences. These CMV-IE sequences are separated by a polylinker sequence, which contains cloning sites unique to the plasmid. Therefore, a cDNA fragment can easily be inserted into the polylinker sequence. Fragments containing ovalbumin or lysozyme cDNA were inserted into this polylinker, as described below and illustrated in figure 4.10.

The first step in pΔTK-Egpt-CMV construction was to delete the *Hind* III and *Bam* HI sites of pΔTK-Egpt, so that the same sites in the polylinker sequence, to be inserted later in the construction, were unique. This was carried out in separate steps, by linearising with the restriction

endonuclease, filling-in with Klenow and then religating. A *Xba* I - *Eco* RI (filled-in) fragment containing the CMV-IE poly(A) signal was removed from the pUC7-CMV cloning plasmid, and then inserted between the *Xba* I and *Nde* I (filled-in) sites of the Δ*Bam HI*, Δ*Hind* III derivative of pΔTK-Egpt. Then, for further simplification, the *Eco* RI site was deleted. This was carried out by *Eco* RI digestion, filling-in with Klenow and then religating, and resulted in the production of pΔTK-Egpt-IE. The most important feature of this plasmid was the presence of a *Sst* I site at the 3'-end of the CMV-IE poly(A) fragment. This allowed the removal of a *Pvu* II - *Sst* I fragment containing the CMV-IE poly(A) signal and some of the IgH-gene sequence, and its replacement with a 0.6Kb *Eco* RI - *Sst* I fragment from a construct called pUC7-CMV-ΔSst (pUC7-CMV cloning plasmid from which the *Sst* I site adjacent to the CMV-IE promoter was deleted, by *Sst* I partial digestion, blunt-ending with T4 DNA-polymerase and then religating). This fragment contains the CMV-IE promoter and poly(A) signal, separated by a polylinker sequence. This insertion resulted in the construction of pΔTK-Egpt-CMV.

A 0.5Kb *Hind* III lysozyme-cDNA fragment (from pTK_2lys) was then inserted into the *Hind* III site of the pΔTK-Egpt-CMV polylinker sequence, and both orientations of the cDNA insert were isolated. $plys^+(d)$ contains the cDNA in the sense orientation with respect to the CMV-IE promoter, and the construct with the alternative orientation of the insert is called $plys^-(d)$ (see figure 4.10.).

A 1.4Kb *Hind* III ovalbumin-cDNA fragment (from pTK_2ov) was similarly inserted into the *Hind* III site of the pΔTK-Egpt-CMV polylinker-sequence. Only the recombinant with the cDNA insert in the sense orientation was isolated. As the MMTV-LTR promoter-element seems to be marginally more active than other promoters in J558L (see sections 3.3 and 3.4.), the CMV-IE promoter was removed by *Eco* RI and *Hind* III digestion (after removal of the *Hind* III site 3' of the ovalbumin-cDNA fragment, by partial *Hind* III digestion, filling-in with Klenow and then religating), and replaced with the *Eco* RI - *Hind* III MMTV-LTR cassette from pUC8-MMTV (see section 3.3.1.), to produce pova(d) (see figure 4.10.).

4.7.2. Cointegration of the non-selectable gene with the gpt gene.

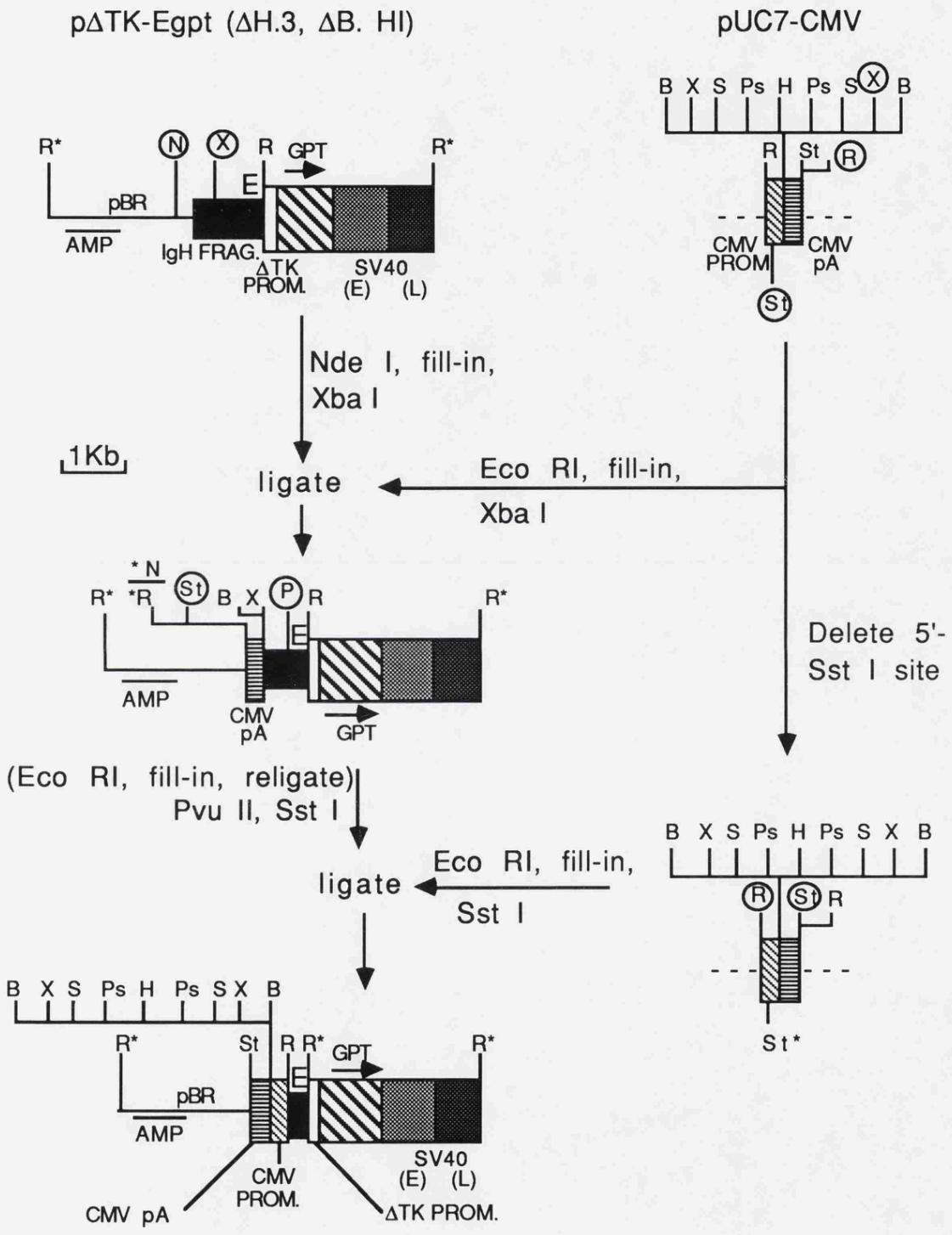
Figure 4.9. Construction of pΔTK-Egpt-CMV.

The strategy used to construct pΔTK-Egpt-CMV is illustrated, and this is described in detail in section 4.7.1. pΔTK-Egpt-CMV is a plasmid which contains *gpt* expression-signals identical to those in pΔTK-Egpt (see section 3.3.1.), except for deleted restriction sites at the junctions between the different fragments. In addition, on the other side of the enhancer, it also possesses fragments containing the promoter and polyadenylation signal derived from the Human CMV-IE-region. These CMV sequences are separated by a polylinker sequence. Sites in this polylinker are unique to the plasmid. Therefore, any coding sequence, from which expression is desired, can easily be inserted into this polylinker sequence.

Briefly, construction of pΔTK-Egpt involved the following manipulations. Firstly, the CMV-IE polyadenylation-signal was inserted upstream of the IgH enhancer in a pΔTK-Egpt derivative, in which some of the restriction sites in the *gpt* transcription unit are deleted. This generated a *Sst* I restriction site, which allowed the CMV-IE polyadenylation-signal fragment to be replaced by a fragment containing both the CMV-IE promoter and the CMV-IE polyadenylation-signal.

The pBR backbone sequences, containing the ampicillin-resistance gene (amp), are represented by thin lines, IgH gene major-intron fragments, containing the enhancer (E), are represented by thin black boxes, the ΔTK promoter is represented by an unfilled box, the *gpt* sequence is represented by a box filled with thick diagonal lines, the SV40 sequences downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the fragment containing the sequence derived from the late region (SV40(L))), the CMV-IE promoter is represented by a box filled with thin diagonal lines and the CMV-IE polyadenylation-signal is represented by a box filled with horizontal lines.

The plasmids illustrated in full are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Nde* I: N, *Xba* I: X, *Eco* RI: R, *Hind* III: H, *Bam* HI: B, *Sal* I: S, *Pst* I: Ps, *Sst* I: St. A polylinker sequence is indicated as PL. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.



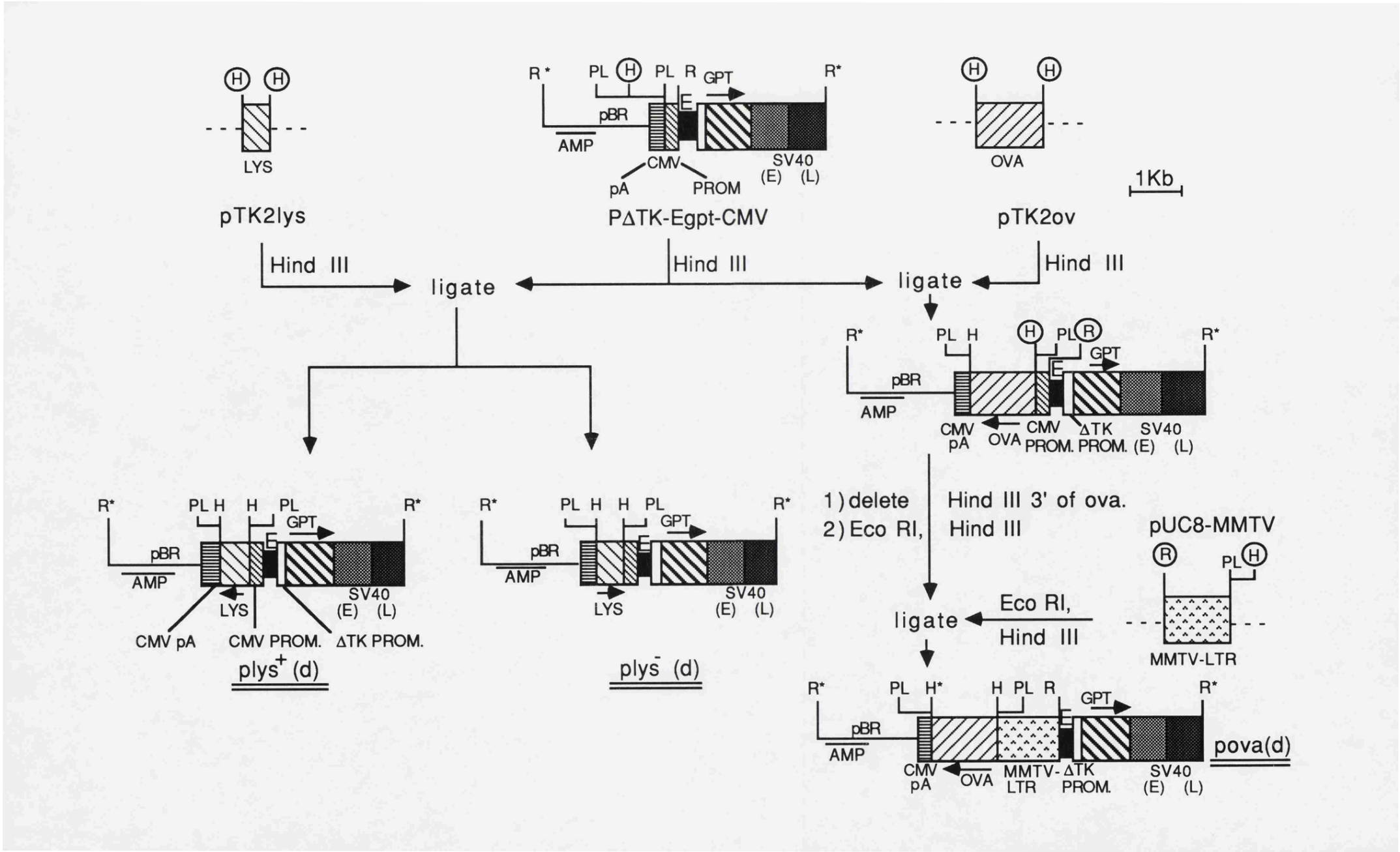
pΔTK-Egpt-CMV

Figure 4.10. Construction of divergent two-gene expression plasmids.

Two-gene expression plasmids, containing divergent transcription units, encoding *gpt*, and ovalbumin or lysozyme, were constructed using the strategy illustrated. As described in section 4.7.1., the ovalbumin or lysozyme cDNA-fragment was simply inserted into the polylinker sequence of pΔTK-Egpt-CMV, which is situated between the CMV-IE promoter and polyadenylation signal. In the case of the ovalbumin expression-plasmid, the CMV-IE promoter was then replaced by the MMTV-LTR. Only the recombinant containing the sense orientation of the ovalbumin-cDNA insert, with respect to the upstream CMV-IE - promoter, was identified, and this is called pova(d). Recombinants containing the lysozyme cDNA in both orientations were identified, and these are called plys$^+$(d) and plys$^-$(d), for the construct containing the sense and missense orientations of the insert, respectively.

The pBR backbone sequences, containing the ampicillin-resistance gene (amp), are represented by thin lines, the IgH gene major-intron fragment, containing the enhancer (E), is represented by a thin black box, the ΔTK promoter is represented by an unfilled box, the *gpt* sequence is represented by a box filled with thick diagonal lines, the SV40 sequences downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequences derived from the late region (SV40(L))), the CMV-IE promoter is represented by a box filled with closely-spaced, thin diagonal lines, the MMTV-LTR is represented by a box filled with triangles, the lysozyme cDNA is represented by a box filled with widely-spaced, thin diagonal lines (extending from top left to bottom right), the ovalbumin cDNA is represented by a box filled with widely-spaced, thin diagonal lines (extending from bottom right to top left) and the CMV-IE polyadenylation-signal is represented by a box filled with horizontal lines.

The plasmids illustrated in full are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. When a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Hind* III: H, *Bam* HI: B, *Xba* I: X, *Sal* I: S, *Pst* I: Ps, *Eco* RI: R. A polylinker sequence is indicated as PL. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.


H
H
LYS
pTK2lys
R*
PL
H
PL
R
GPT
R*
pBR
AMP
CMV
pA
PROM
SV40
(E)
(L)
PΔTK-Egpt-CMV
H
H
OVA
pTK2ov
1Kb
Hind III
ligate
Hind III
ligate
Hind III
R*
PL H
H
PL
R
GPT
R*
pBR
AMP
CMV pA
OVA
CMV PROM.
ΔTK PROM.
SV40
(E)
(L)
1) delete
2) Eco RI,
Hind III 3' of ova.
Hind III
pUC8-MMTV
R
PL
H
MMTV-LTR
Eco RI,
Hind III
ligate
R*
PL H
H
PL
GPT
R*
pBR
AMP
LYS
SV40
(E)
(L)
CMV pA
CMV PROM.
ΔTK PROM.
plys+ (d)
R*
PL H
H
PL
GPT
R*
pBR
AMP
LYS
SV40
(E)
(L)
plys- (d)
R*
PL
H*
H
PL
R
GPT
R*
pBR
AMP
CMV pA
OVA
MMTV-LTR
ΔTK PROM.
SV40
(E)
(L)
pova(d)

J558L was transfected with pova(d) using spheroplast fusion, and resultant, stable *gpt*$^{+}$-transfectant clones were isolated and expanded. Before screening transfectants for expression of ovalbumin, it was important to confirm cointegration of the ovalbumin-cDNA with the *gpt* gene had occurred. The results of section 4.6. seemed to suggest that cointegration of μ and *gpt* genes occurs at a high frequency, following transfection of J558L with the two-gene (μ and *gpt*) plasmid, pSV-Vμ1. However, the spheroplast-fusion transfection with pSV-Vμ1 was performed using the *E. coli* strain, DH5. In contrast, the simple cotransfections described in section 4.2.3. (and all other transfections) were performed using the *E. coli* strain, TG1. In addition, pSV-Vμ1 differs greatly in structure, from the plasmids used in the simple cotransfections and also pova(d). Although unlikely, it could not be discounted that host-strain or plasmid-structure considerations, rather than the physical linkage of the μ and *gpt* genes in pSV-Vμ1, contributed to their high stable-cointegration frequency. Also, it had not been demonstrated that the two-gene plasmids become integrated into the host genomic-DNA. Therefore Southern-blotting analysis was performed.

The Southern blot in figure 4.11. shows that the ovalbumin-cDNA was cointegrated with the *gpt* gene in all the pova(d) transfectants analysed. The blot contained *Hind* III / *Bam* HI genomic digests, and the autoradiogram shows that the ovalbumin-cDNA fragment was released as the expected 1.4Kb fragment, from the ≈7Kb pova(d)-backbone which contains the *gpt*-SV40 sequence. The copy number of the integrated plasmid sequences varied from clone to clone, but was generally high and all clones integrated more than one plasmid. It appears that each clone contained the *gpt* and ovalbumin sequences at equivalent copy number (allowing for a degree of error in quantitations in which signal intensities in control tracks are used to determine copy numbers of fragments in the genomic tracks). In most transfectants, the observations are consistent with a single concatameric-unit, containing multiple units of transfected plasmid, arranged as head-to-tail tandem repeats, having been integrated into the genome.

4.7.3. Production of ovalbumin and lysozyme transcripts from transfected genes.

Figure 4.11. Southern-blot analysis of pova(d) transfectants of J558L.

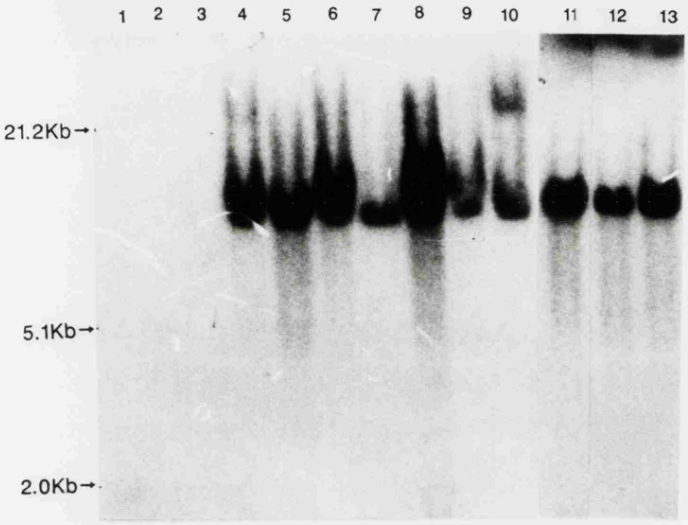
Ten monoclonal pova(d)-transfectants of J558L were expanded, and then analysed for cointegration of the *gpt* and ovalbumin-cDNA sequences. Genomic DNA from each transfectant was digested with *Hind* III and *Bam* HI, run out on a 1% agarose gel and Southern blotted (see section 2.11.3.). The blot was first hybridised to the *gpt*-SV40 probe (autoradiogram A), then stripped and hybridised to the ovalbumin probe (autoradiogram B). The probes are described in section 2.10.1.

The tracks of the blot shown contained the following samples: track 1: *Hind* III / *Bam* HI digests, of 10μg untransfected-J558L genomic DNA, and of pSV_2gpt, containing 1.5pg of the 1.9Kb *Hind* III - *Bam* HI *gpt*-SV40 fragment (a positive control); track 2: *Hind* III / *Bam* HI digests, of 10μg untransfected-J558L genomic DNA, and of a plasmid, pSV_2ov, containing 1.5pg of the 1.4Kb *Hind* III ovalbumin-cDNA fragment (a positive control); track 3: 10μg *Hind* III / *Bam* HI digested untransfected-J558L genomic DNA (negative control); tracks 4-13: genomic digests of the following pova(d) transfectants: track 4: clone D7; track 5: clone B7; track 6: clone C5; track 7: clone B11; track 8: clone B10; track 9: clone E9; track 10: clone E6; track 11: clone D7; track 12: clone C4; track 13: clone D9.

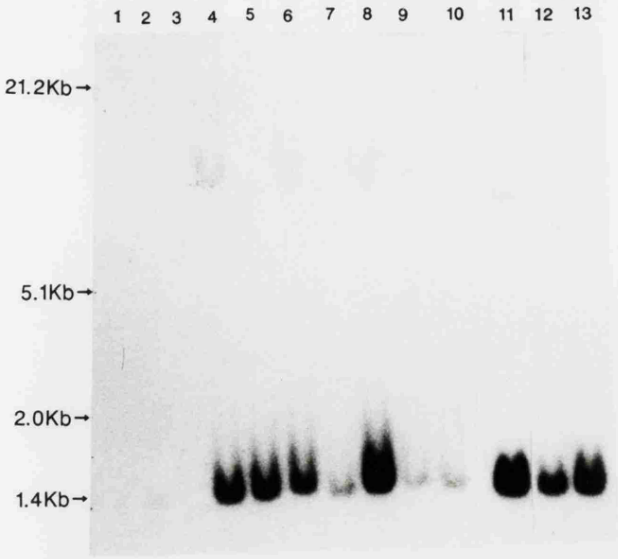
Autoradiogram A shows that the *gpt*-SV40 probe hybridises to the ≈ 7Kb *Hind* III / *Bam* HI backbone fragment of pova(d), which contains the complementary *gpt*-SV40 sequence (see figure 4.10.). It also hybridises to the same 1.9Kb *gpt*-SV40 fragment from pSV_2gpt, in track 1 (just visible on the photograph).

Autoradiogram B shows that the ovalbumin probe hybridises to the 1.4Kb ovalbumin-cDNA fragment, released from integrated pova(d)-sequences by the *Hind* III / *Bam* HI digest. It also hybridises to the 1.4Kb *Hind* III ovalbumin-cDNA fragment from pSV_2ov, in track 2 (again just visible on the photograph but an even lower-intensity signal than from the same amount (1.5pg) of control *gpt*-SV40 fragment in autoradiogram A). No hybridisation is visible to track 3 in either autoradiogram, illustrating the specificity of the hybridising genomic-DNA fragments to transfected cells.

In summary, for the ten pova(d) transfectants analysed, a 100% cointegration frequency of the non-selected ovalbumin-cDNA sequence with the selected *gpt*-sequence is observed. The copy number of the integrated plasmids is generally high (same calculation of copy number applies as explained for figure 4.3.), but it is highly variable between clones. However, the same copy-number of *gpt* and ovalbumin sequences appear to be present in each clone. These results are discussed in detail in section 4.7.3.



AUTORADIOGRAM A
(gpt-SV40 probe)



AUTORADIOGRAM B
(ovalbumin probe)

pova(d), plys$^+$(d) and plys$^-$(d) stable-transfectants of J558L were analysed for transcription of the non-selected gene by RNA-dot-blot analysis (see section 2.12.5.). Dots of total RNA were hybridised to the appropriate RNA-complementary single-stranded DNA probe (see section 2.10.1.). Figure 4.12. shows that three pova(d) transfectants, which contained integrated ovalbumin-cDNA (as shown in figure 4.11.), were transcribing the ovalbumin coding-sequence. Two out of four monoclonal plys$^+$(d)-transfectants (not previously checked for cointegration of the lysozyme-cDNA with the *gpt* gene) were transcribing the lysozyme sequence at detectable levels. Interestingly, it appears that the lysozyme sequence was also being transcribed, albeit at a lower level, in two of the plys$^-$(d) transfectants analysed. This is surprising, because in this plasmid the lysozyme coding-sequence is in the wrong orientation for transcripts to be initiated from the CMV-IE promoter. This observation is discussed in greater detail in section 5.2.2. It must be emphasised that dot-blot analysis only indicates the presence of transcripts. Northern analysis and RNA-mapping analysis is required to show that mature mRNA with the expected structure is being produced. However, the above result was sufficient to justify an analysis for expression of the encoded proteins, and only if expression was not obtained would more detailed RNA analysis be necessary.

4.7.4. Analysis for lysozyme and ovalbumin expression.

The proteins of transfectant cell-lines, shown above to be positive for heterologous-RNA expression, were biosynthetically-labelled with ^{35}S-methionine. Supernatants were then analysed for secretion of lysozyme or ovalbumin as appropriate, by immune-precipitation from cell supernatants with rabbit anti - chicken-lysozyme antibody (see section 2.4.1.) and commercially supplied rabbit anti - chicken-ovalbumin antibody respectively.

Figure 4.13. demonstrates that both of the plys$^+$(d) transfectants analysed, clones E8 and F3, secreted lysozyme (assuming it was not released into the medium by cell lysis). In contrast, the plys$^-$(d) transfectant, clone C5, which generated lysozyme transcripts, did not express detectable levels of lysozyme. This might be because the transcripts are unstable and unable to form stable and/or mature mRNA. If,

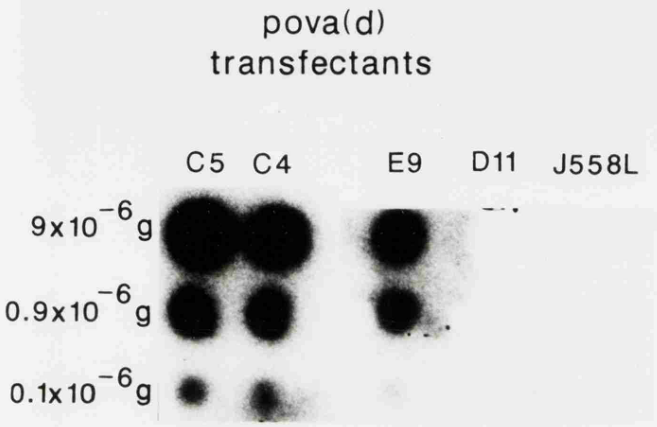
Figure 4.12. RNA dot-blot analysis to test for transcription of the non-selectable gene in transfectants containing two-gene expression plasmids.

Following the demonstration that transfection of J558L with pova(d) leads to a high cointegration-frequency of *gpt* and ovalbumin-cDNA sequences (see figure 4.11.), it was then necessary to ensure that the ovalbumin-cDNA sequence is transcribed. It was expected that plys$^+$ (d) and plys$^-$(d) transfectants would also give a high cointegration frequency of *gpt* and lysozyme-cDNA sequences, because of the similar structure of the plasmids to pova(d). Hence, lysozyme RNA production was analysed for in such transfectants, without first analysing by Southern blotting.

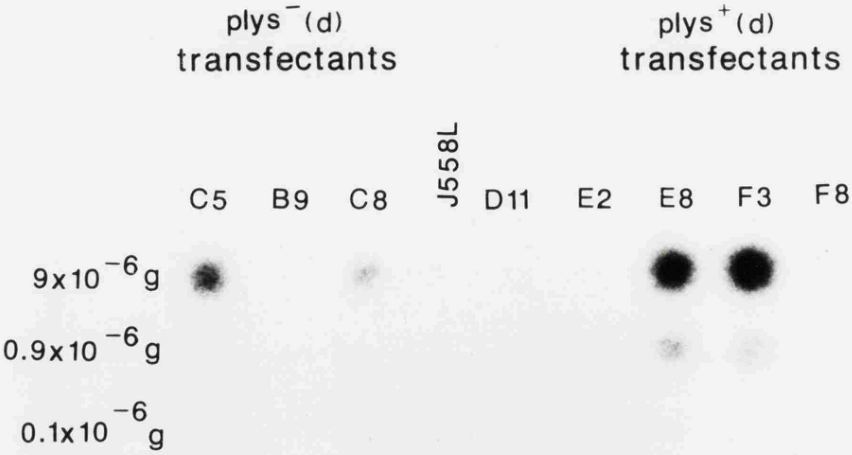
Total RNA was isolated from transfectant clones (see section 2.12.2.), and serial dilutions were made and dotted onto Hybond-N filters (see section 2.12.5.). Then the RNA was hybridised to RNA-strand-complementary single-stranded DNA probes for ovalbumin or lysozyme (see section 2.10.1.) as appropriate.

Autoradiogram A shows the hybridisation of the ovalbumin probe to RNA from three pova(d) transfectants (clones C5, C4 and E9). RNA samples from untransfected J558L and a pCMV-Egpt (see section 3.3) transfectant of J558L, clone D11, act as negative controls. The autoradiogram suggests that all clones generate ovalbumin transcripts. Figure 4.11. showed that all these pova(d) transfectants contain integrated ovalbumin-cDNA sequences, but at different copy numbers. There is a loose correlation between gene copy-number and RNA levels, because clone C5, which contains the highest copy-number, gives higher levels of ovalbumin RNA than clone E9, which contains the lowest copy-number.

Autoradiogram B shows the hybridisation of the lysozyme probe to RNA from plys$^+$(d) and plys$^-$(d) transfectants. Clones E2, E8, F3 and F8 are plys$^+$(d) transfectants, and it can be seen that lysozyme transcripts are detectable in clones E8 and F3. Interestingly, two of the three plys$^-$(d) transfectants analysed, also give a positive signal for lysozyme transcripts (above the background level from the negative control samples, which are the same as described for autoradiogram A). However, the signal intensity for RNA from plys$^-$(d) transfectants is low, compared with that from the plys$^+$(d) transfectants which generate lysozyme transcripts.



AUTORAD. A (ovalbumin probe)



AUTORAD. B (lysozyme probe)

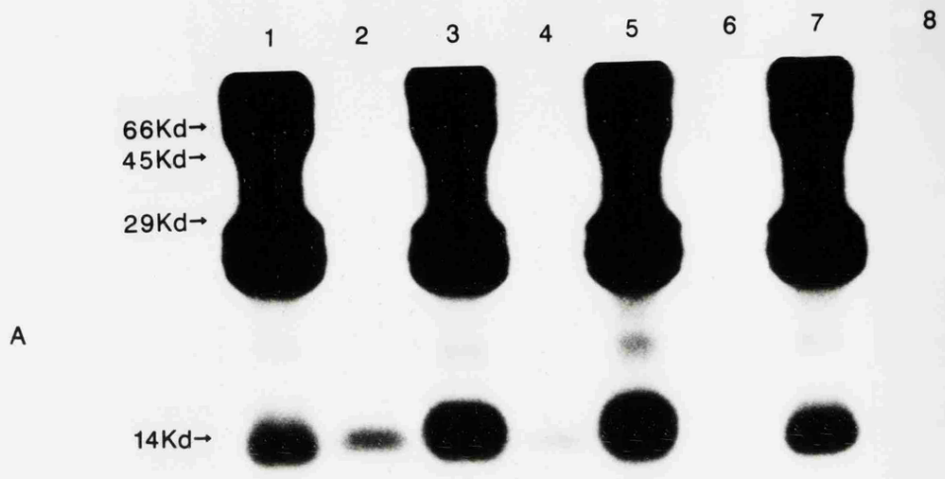
Figure 4.13. Analysis of lysozyme and ovalbumin RNA-expressing transfectants for secretion of the encoded protein.

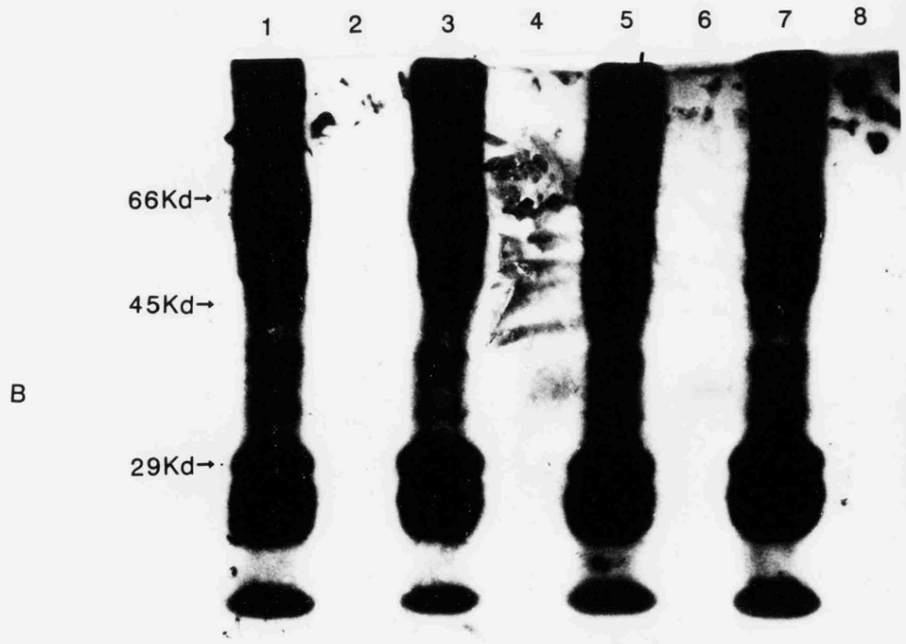
plys$^+$(d), plys$^-$(d) and pova(d) transfectants of J558L, which were shown to generate transcripts from the non-selected gene (see figure 4.12.), were incubated with ^{35}S-methionine in order to biosynthetically-label protein (see section 2.5.1.). Secreted lysozyme or ovalbumin was immune-precipitated from cell supernatants using the appropriate antibody, and analysed by SDS-PAGE and autoradiography (see sections 2.5.2. and 2.5.3. respectively).

A: an autoradiogram of a 13% SDS-PAGE gel, containing crude cell-supernatants, and rabbit anti - chicken-lysozyme antibody immune-precipitates from cell supernatants (all from equal volumes of supernatants). The samples in each track and the cell line from which they were obtained are as follows: track 1: untransfected J558L crude-supernatant; track 2: plys$^+$(d) transfectant (clone F3) immune-precipitate; track 3: plys$^+$(d) transfectant (clone F3) crude supernatant; track 4: plys$^+$(d) transfectant (clone E8) immune-precipitate; track 5: plys$^+$(d) transfectant (clone E8) crude supernatant; track 6: plys$^-$(d) transfectant (clone C5) immune-precipitate; track 7: plys$^-$(d) transfectant (clone C5) crude supernatant; track 8: untransfected J558L 'immune-precipitate' (negative control).

B: an autoradiogram of an 11% SDS-PAGE gel, containing crude cell-supernatants, and rabbit anti - chicken-ovalbumin antibody immune-precipitates from cell supernatants (all from equal volumes of supernatants). The samples in each track and the cell line from which they were obtained are as follows: track 1: pova(d) transfectant (clone E9) crude supernatant; track 2: pova(d) transfectant (clone E9) immune-precipitate; track 3: pova(d) transfectant (clone C5) crude supernatant; track 4: pova(d) transfectant (clone C5) immune-precipitate; track 5: pova(d) transfectant (clone C4) crude supernatant; track 6: pova(d) transfectant (clone C4) immune-precipitate; track 7: untransfected-J558L crude supernatant; track 8: untransfected J558L 'immune-precipitate' (negative control).

In autoradiogram A, the immune-precipitate tracks, tracks 2 and 4, show that the 14KD protein, lysozyme, is present in the supernatant from the two plys$^+$(d) transfectants, clones F3 and E8. As expected, the plys$^-$(d) transfectant, C5 (which generates lysozyme 'transcripts'; see figure 4.12.), does not secrete detectable levels of lysozyme. In autoradiogram B, ovalbumin secretion from the pova(d) transfectants is not convincingly detectable. If molar levels of expression of lysozyme and ovalbumin are the same, the signal, relative to other bands, from the lysozyme band in autoradiogram A should only be about twice as intense as the signal from the ovalbumin band in autoradiogram B. This is because ovalbumin contains three methionine residues compared with two for lysozyme, and lysozyme and ovalbumin have molecular weights of 14KD and 45KD respectively. However, no bands corresponding to the size of ovalbumin are visible in the immune-precipitate tracks. In the immune-precipitate tracks of clones C5 and C4, a faint band corresponding to a molecular weight of approximately 35KD is present. This is unlikely to be a degraded derivative of ovalbumin because it is also present, at even lower levels (not visible on photograph), in the untransfected J558L 'immune-precipitate' (track 8), suggesting it is non-specific. Its presence is possibly due to some inconsistent event such as cell lysis.


1
2
3
4
5
6
7
8
66Kd→
45Kd→
29Kd→
A
14Kd→


1
2
3
4
5
6
7
8
66Kd→
45Kd→
B
29Kd→

as is proposed in section 5.2.2., the transcripts generated are produced by read-through of RNA-polymerase from the *gpt* transcription-unit, the 3'-processing of the *gpt* transcript might generate a lysozyme-coding transcript without a 5'-cap site. The cap site is normally recognised as part of the translation-initiation signal and may provide protection against exonucleases (see section 1.9.2.). Also, and no matter where the origin of the transcripts, there is no recognised polyadenylation-signal available immediately downstream of the lysozyme coding-sequence. As discussed in section 1.9.2., polyadenylation is also essential for the production of stable mRNA. When no site is available, 3'-processing might occur around cryptic poly(A)-signals, but these are likely to be less efficient than *bone fide* poly(A) signals.

Interestingly, the lysozyme seems to comigrate during SDS-PAGE analysis with a polypeptide produced in all J558L cells (see figure 4.13.). This may be a result of cell lysis, but its high level (assuming its methionine content is typical of total protein) relative to other proteins (except the ≈ 25KD λ-chain) suggests it may be a secreted protein. One possibility is that it is a shortened version of an Ig chain. Such proteins have been identified as being secreted by numerous myeloma cell-lines (Scharff *et al.*, 1970; Birshtein *et al.*, 1974).

Figure 4.13. also shows, in contrast, that no ovalbumin secretion could be convincingly detected from three pova(d)-transfectants (clones C4, C5 and C9). Ideally, more transfectants would have been screened to increase the statistical significance of this result. If ovalbumin secretion could not be detected, there are many potential reasons why this might be so, as outlined below.

One possibility is that the linearisation of the plasmid, prior to integration, occurred within the ovalbumin coding-sequence. The Southern blot shown in figure 4.11. indicates that the plasmids were integrated as a concatameric unit and, therefore, it might be expected that the linking of plasmids in a tandem head-to-tail fashion could reform cleaved sequences. Indeed, at the limited level of resolution on the Southern blot, it appears that all ovalbumin sequences, in the transfectants analysed, were preserved as the complete 1.4Kb cDNA-fragment. However, reports suggest that homologous recombination, which is believed to be a process by which the concatameric units of plasmid form, prior to their chromosomal integration

(see section 1.6.), is primarily non-conservative (Seidman, 1987). Hence, if homologous recombination occurs within the ovalbumin coding-sequence during formation of the concatameric unit, the reading frame may be changed resulting in the production of a missense protein, or a translation-termination codon may be introduced resulting in the production of a truncated protein. However, if the site of plasmid linearisation is random, only about a quarter (1.4Kb accounts for about 1/4 of the size of pova(d)) of pova(d) transfectants should experience this problem. Therefore, it is unlikely that this is an explanation for non-expression in all three pova(d)-transfectants.

Detailed RNA analysis (Northern, RNA-mapping, nuclear run-off) is required to show that mRNA of an expected size and structure is being produced, and also to quantify the level of mature mRNA, so as to test if the low expression is due to a site unfavourable for transcription or because of RNA structure. The presence of the CMV-IE polyadenylation signal downstream of the ovalbumin-cDNA fragment (see figure 4.10.) should ensure 3'-processing of ovalbumin transcripts, as it presumably does on lysozyme transcripts to allow for lysozyme expression. However, the promoter upstream of the ovalbumin-cDNA in pova(d) is contained in a MMTV-LTR sequence, whereas a CMV-IE promoter is upstream of the lysozyme-cDNA fragment in plys$^+$(d). It might be suggested that the CMV-IE promoter is more appropriate than the MMTV-LTR, perhaps because it drives transcription at a higher frequency, or the contribution of this fragment to the 5'-untranslated region of the RNA contributes to RNA stability. This is unlikely to be so, however, because expression from the *gpt* gene in pMMTV-Egpt and pMMTV-gpt transfectants seemed to be at least as efficient as that from transfectants containing the corresponding CMV-IE-promoter - containing constructs, as seen from the stable *gpt*$^+$-transfection frequencies obtained following their introduction into J558L (see tables 3.1. and 3.2.). Even so, because the coding sequences (ovalbumin and *gpt*) are different, it is not possible to exclude the possibility that a MMTV-LTR - derived 5'-untranslated region in an ovalbumin transcript forms a secondary structure, which is detrimental to RNA stability or translation. As discussed in section 1.9., the expression of any gene product from any construct has to be determined empirically. A cell-free translation experiment could be used to confirm that the mRNA can be translated (*in*

vitro !) and encodes a complete ovalbumin protein-sequence.

I believe the most likely reason for non-expression of ovalbumin from pova(d) transfectants is at the post-translational level. For example, ovalbumin may not be secreted efficiently from J558L. One reason for this could be that the untypical, internal signal-sequence of ovalbumin (see section 4.1.) is not recognised by J558L, preventing its entry into the secretory pathway. Alternatively, ovalbumin may not pass through the secretory pathway efficiently. In fact, Colman and Ceriotti (1988) reported that ovalbumin is secreted from *Xenopus* oocytes much less efficiently than lysozyme, and accumulates at much higher levels intracellularly. However, in preliminary analysis of radiolabelled-cell lysates, I could not detect intracellular accumulation of ovalbumin in the three pova(d)-transfectants, using immune-precipitation. In contrast, in similar experiments, μ chain and lysozyme were detected in cell lysates of pSV-Vμ1 and plys$^+$(d) transfectants respectively (not shown). Perhaps, if ovalbumin does not enter the secretory pathway, it is rapidly degraded, and this would prevent its intracellular detection.

Unfortunately, there was no convenient, available positive-control for confirming that the system for detecting ovalbumin in cell supernatants by immune-precipitation was functional; ideally, an ovalbumin-secreting cell line would be used. However, in Ouchterlony immunodiffusion tests (see section 2.4.1.), 10μl of a 1:9 (v/v) dilution of the commercially supplied rabbit anti - chicken-ovalbumin antiserum, in NET buffer, detected ovalbumin in 10μl of a 10μg/ml solution, but it required 10μl of undiluted rabbit anti - chicken-lysozyme antiserum to detect the same amount of lysozyme (not shown). This suggested the former antiserum had the higher titre. Also, as both antisera were raised in rabbits, the antibodies are likely to have similar affinities for the protein-A on *S. aureus* membranes (see sections 2.4.2. and 2.5.2.) and, therefore, behave similarly in the immune-precipitation procedure..

Several possible explanations of why ovalbumin expression was not detected from pova(d) transfectants have been listed above. There could also be more trivial explanations, such as a mutation being introduced into the ovalbumin coding-sequence during DNA-subcloning manipulations. Therefore, rather than exhaustively clarifying the reasons for the apparent non-expression of ovalbumin, efforts were concentrated on lysozyme

expression; the secretion of lysozyme was used as a model system on which attempts were made to increase the secretion (expression) level. Indeed, at this point in the study, the cell-line secreting the highest level of lysozyme was the plys$^+$(d) transfectant, clone F3 (see figure 4.13.). Its level of lysozyme secretion is more than an order of magnitude lower than IgG expression from X63.Ag8, in molar terms (as indicated in the legend associated with figure 4.15.); it was the aim to secrete lysozyme at least at similar levels.

4.7.5. Variation of the polyadenylation-signal 3' of the lysozyme cDNA.

For most eucaryotic genes, cleavage and polyadenylation at the 3'-end of the primary transcript is essential for stable mRNA formation (see section 1.9.2.). Therefore, if the efficiency of a heterologous polyadenylation-signal is not optimal, it might limit the expression levels from a gene. Also, a heterologous fragment containing a polyadenylation signal contributes part of the non-translated 3'-end of the mRNA, which may influence the stability of the mRNA. For these reasons, the CMV-IE polyadenylation signal in plys$^+$(d) was replaced by a series of other available fragments containing a polyadenylation signal. It was expected that if any polyadenylation signal was particularly compatible with J558L, this would be reflected in a high level of lysozyme expression, in cells transfected with plasmids containing this polyadenylation-signal downstream of the lysozyme-cDNA fragment.

plys$^+$(d)-derivative lysozyme expression-plasmids were constructed which contained one of the following fragments, 3' of the lysozyme cDNA; the SV40, early poly(A)-signal (plys-SV40); the SV40, early poly(A)-signal with the small t-antigen intron upstream (plys-ISV40) [this is the same fragment as downstream of the *gpt* transcription unit]; the HSV - thymidine-kinase poly(A) signal (plys-TK); and the rabbit - β-globin poly(A) signal (plys-βglo). In all other respects the plasmids are identical to plys$^+$(d). Multiple subcloning procedures were required to allow the construction of the plasmids and these details are described in the figure legend of figure 4.14.; this figure illustrates the structure of the plasmids.

I transfected the above series of plasmids into J558L and

Figure 4.14. plys⁺(d) derivatives containing different polyadenylation signals.

The CMV-IE polyadenylation-signal, downstream of the lysozyme-cDNA in plys⁺(d), was replaced with other polyadenylation signals, as described below (also, see section 4.7.5.), to produce the illustrated plasmids. Other than the different polyadenylation-signals, all plasmids have identical structures and, therefore, only restriction sites flanking the fragments containing these polyadenylation signals are illustrated.

The pBR backbone, containing the ampicillin-resistance gene (amp), is represented by a thin line, the ΔTK promoter is represented by an unfilled box, the *gpt* sequence is represented by a box filled with thick diagonal lines, SV40 sequences downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequence derived from the late region (SV40(L))), the IgH gene major-intron fragment, containing the enhancer (E), is represented by a thin black box, the CMV-IE promoter is represented by a box filled with closely-spaced, thin diagonal lines, and the lysozyme cDNA is represented by a box filled with widely-spaced, thin diagonal lines. The polyadenylation-signal - containing fragments are represented by boxes filled with the following patterns: CMV-IE (horizontal lines), SV40 early, with (SV40(E)) or without (SV40(E)pA) the small t-antigen intron, (darkly shaded as for the same sequence(s) immediately downstream of the *gpt* transcription-unit), TK (broken vertical lines) and β-globin (broken horizontal lines).

The plasmids are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Restriction sites are abbreviated as follows: *Bam* HI: B, *Xba* I: X, *Sal* I: S, *Pst* I: Ps, *Hind* III: H, *Sst* I: St, *Nde* I: N, *Kpn* I: K. Sites deleted during construction of the plasmids are indicated with asterisks.

Plasmid construction.

The details below describe the subcloning manipulations performed in order to make the plasmids illustrated in this figure.

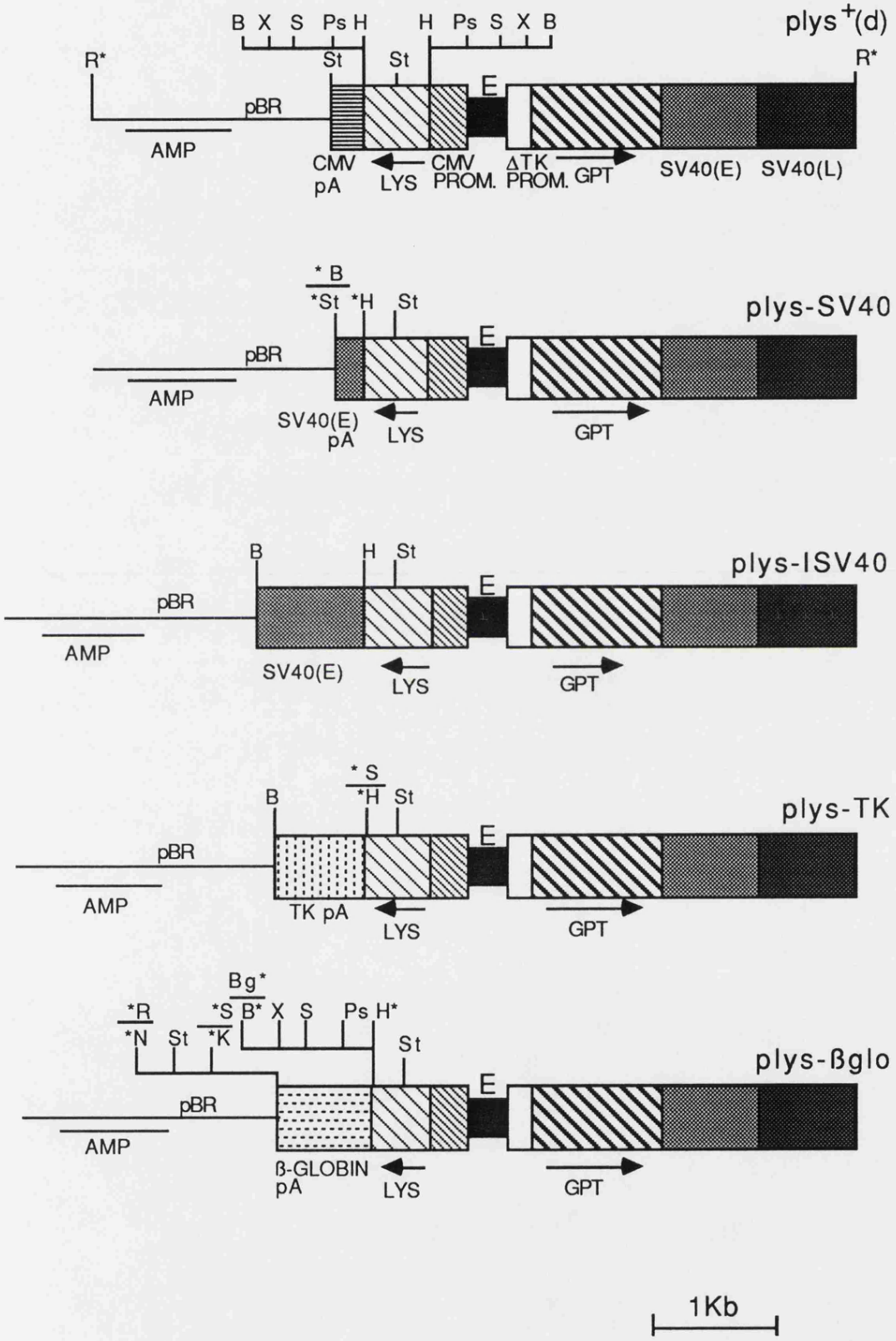
The *Hind* III site at the 3'-end of a lysozyme-cDNA fragment in the plasmid, pTK_2lys, was deleted, by partial *Hind* III digestion, filling-in with Klenow and then religating. This allowed a 1.6Kb *Hind* III - *Bam* HI (filled-in) fragment, containing the lysozyme cDNA and the SV40, early polyadenylation-signal downstream, to be removed, and this was inserted into the *Hind* III and *Sst* I (cut-back) sites of the plys⁺(d) backbone. Identification of the required recombinant was facilitated by regeneration of the *Bam* HI site from fusion of the blunt-ended *Bam* HI and *Sst* I sites. The net result of these manipulations was substitution of the SV40, early polyadenylation-signal for the CMV-IE polyadenylation-signal in plys⁺(d), to produce plys-SV40.

A plasmid, pTK_1lys, was digested with *Sst* I (which cut within the lysozyme coding-sequence) and *Bam* HI, to release a fragment containing the 3'-end of the lysozyme cDNA and downstream SV40 sequences. The SV40 fragment contains the small t-antigen intron and the early polyadenylation-signal, and is identical to the SV40 fragment downstream of the *gpt* sequence in many of the plasmids described in this study (e.g. plys⁺(d)). The *Bam* HI site at the 5'-end of the lysozyme-cDNA fragment in plys-SV40

Figure 4.14. (cont.)
was deleted, by partial *Bam* HI digestion, filling-in with Klenow and then religating. This allowed a *Sst* I - *Bam* HI fragment, containing the 3'-end of the lysozyme cDNA and the SV40, early polyadenylation-signal, to be removed. Then the above described *Sst* I - *Bam* HI fragment was inserted into the backbone, to produce plys-ISV40.

A 1.9Kb *Eco* RI - *Pvu* II fragment, containing most of the HSV-TK coding-sequence and also the polyadenylation signal, was isolated (from the plasmid, pTK1; this is not related to pTK_1lys which was used in plys-TK construction) and then cloned into the *Eco* RI and *Sma* I sites of the pUC18 polylinker. A 0.7Kb *Sma* I - *Bam* HI fragment containing the TK polyadenylation-signal was then removed from the insert, and cloned into the *Hind* III (filled-in) and *Bam* HI sites of pTK_2lys-Δ*Hind* III (pTK_2lys from which the *Hind* III site upstream of the lysozyme cDNA was deleted, by partial *Hind* III - digestion, filling-in and then religating). This generated another *Sst* I - *Bam* HI cassette, which contains the 3'-end of the lysozyme cDNA and the TK polyadenylation-signal. This was inserted into the Δ*Bam* HI derivative of plys-SV40, in the same manner as for plys-ISV40 construction, to produce plys-TK.

A 0.7Kb *Sal* I (filled-in) - *Bgl* II fragment (from pβ5'SV-BglII), containing the rabbit β-globin polyadenylation-signal, was cloned into the *Bam* HI (compatible with *Bgl* II) and *Kpn* I (cut-back) sites of the pUC18 polylinker. In a recombinant isolated, the *Sal* I site was not regenerated as expected from the fusion of blunt-ended *Sal* I and *Kpn* I sites, but the size of the insert was indiscernable (on a 1% agarose gel) from that of the fragment cloned into the polylinker. As the polyadenylation signal sequences are not close to this junction, it was considered unlikely that this subcloning step had damaged the poyadenylation signal. The insert was removed from the pUC polylinker on an *Xba* I - *Eco* RI fragment, and cloned between an *Xba* I site (in the polylinker sequence) and the *Eco* RI site (downstream of the CMV-IE polyadenylation signal) in plysE(t) (see figure 4.20.). From the resultant plasmid, a *Sst* I fragment, containing the 3'-end of the lysozyme cDNA and the β-globin polyadenylation-signal downstream, was removed, and then cloned into the *Sst* I sites of $plys^+$(d). A recombinant containing the insert in the sense orientation (i.e. producing a complete lysozyme-cDNA sequence) was identified. This was named plys-βglo.


plys+(d)
B X S Ps H
H Ps S X B
R*
St
St
R*
E
pBR
AMP
CMV pA
LYS
CMV PROM.
ΔTK PROM.
GPT
SV40(E)
SV40(L)
plys-SV40
* B
*St *H St
E
pBR
AMP
SV40(E) pA
LYS
GPT
plys-ISV40
B
H St
E
pBR
AMP
SV40(E)
LYS
GPT
plys-TK
* S
B
*H St
E
pBR
AMP
TK pA
LYS
GPT
plys-ßglo
Bg*
*R
S B X S Ps H*
*N St *K
St
E
pBR
AMP
ß-GLOBIN pA
LYS
GPT
1Kb

selected for stable *gpt*$^+$-transfectants. I planned to compare the relative levels of lysozyme expression from the constructs, by measuring lysozyme secretion from polyclonal populations containing hundreds of transfectant clones, assuming that they give statistically-representative levels of expression. Unfortunately, the number of stable *gpt*$^+$-transfectants generated from these divergent constructs was surprisingly low (chapter 5 is devoted to attempts to explain this) and, therefore, sufficient clones were not available (≈ 1-20 clones per transfection experiment) to allow analysis of representative transfected-cell populations. Consequently, the method of analysis used was much less satisfactory; about six stable *gpt*$^+$-transfectant clones produced with each plasmid were expanded, and comparative levels of lysozyme secretion were analysed, using the biosynthetic-radiolabelling / immune-precipitation method. In the clones analysed, lysozyme secretion was very weak (on SDS-PAGE gels) from those containing the SV40 early (with intron), TK and β-globin downstream-sequences (not shown), whereas two high-level expressing transfectants, one containing plys$^+$(d), and the other containing plys-SV40, were identified (their analysis is described in section 4.7.6.). However, it is impossible to conclude that these plasmids contain polyadenylation signals more suitable for expression of lysozyme in J558L; more extensive screening may have allowed identification of high-level expressing clones from transfectants containing any of the plasmids. In fact, preliminary analysis of transiently introduced plasmids suggested that plys-βglo gives the highest levels of lysozyme expression. As discussed in section 1.9.5., globin mRNA is particularly stable, and this may be due in part to the non-translated 3'-region. Such a sequence might also allow for increased stability of lysozyme transcripts. However, if this is so, it is surprising that none of the stable plys-βglo - transfectants tested expressed high levels of lysozyme.

4.7.6. Study of transfectants secreting high levels of lysozyme.

In the above described analysis on SDS-PAGE gels for comparison of the effect of different polyadenylation signals on lysozyme expression, two J558L transfectants, clones D8 and C6, apparently secreting high levels of lysozyme, were identified. These were isolated for further studies.

Clone D8 is a plys$^+$(d) transfectant, and clone C6 is a plys-SV40 transfectant.

Interestingly, both cell lines were obtained from clones which had been growing slowly in XHMPA selective-medium during selection, and which had to be removed from this medium to allow their rapid expansion. Lysozyme expression was purposefully measured from cell lines derived from both fast and slow growing colonies in order to test if there was any difference in their levels of lysozyme expression. It was predicted that the fast growing colonies would express lysozyme at higher levels. This was because it was hypothesised that high-level expression from the *gpt* gene (presumably reflected by vigorous growth in XHMPA selective-medium) reflects a site of integration favourable for high-level expression or, alternatively, reflects a high plasmid copy number; both of these scenarios might also ensure high-level expression from the cointegrated non-selected gene. The fact that the opposite result was obtained may be statistically insignificant. Alternatively, it may be due to preferential expression from one of the two transcription units (chapter 5 is devoted to a study of this potential problem), or because high-level lysozyme expression is detrimental to the cloning efficiency of a cell due to a drain on its resources (see chapter 6).

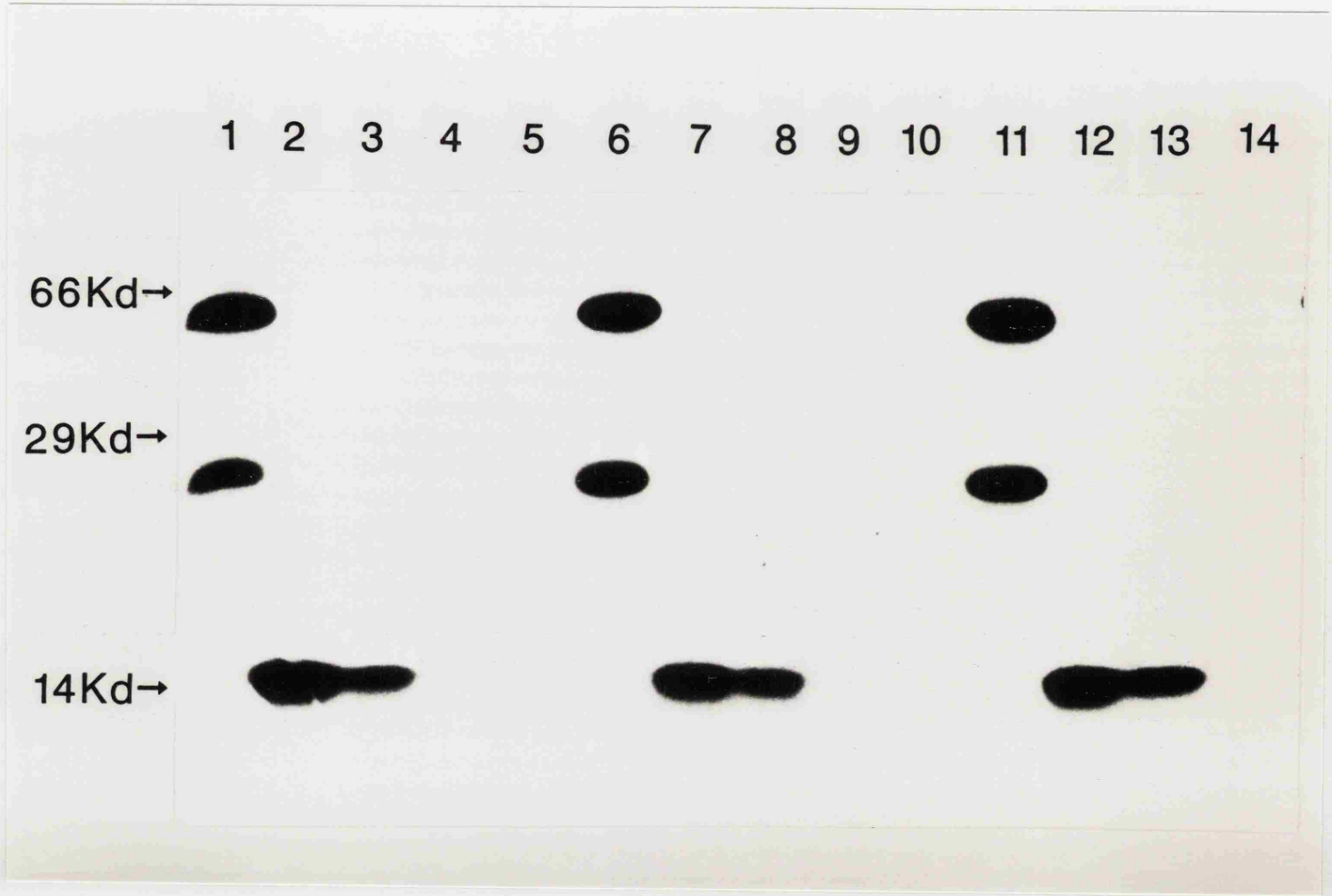
An estimation of the lysozyme secretion-level from clones D8 and C6 was made by comparison with the level of secretion of $IgG_1(\kappa)$ from X63.Ag8, using the biosyntheic-radiolabelling / immune-precipitation assay. In order to obtain reproducible expression-level ratios, it was important to maintain cells in optimal condition. Figure 4.15. shows a SDS-PAGE gel containing immune-precipitates of secreted proteins, from three separate sets of radiolabelling incubations, which were performed at 48-hour intervals. It illustrates the reproducibility of the quantitation and shows that D8 is the highest-level lysozyme-secreting transfectant. Up to this point, the highest level of lysozyme secretion was from the plys$^+$(d) transfectant, clone F3 (see section 4.7.4.). At the exposure of the photograph shown, the lysozyme secreted from this clone is not visible, but comparison of band intensities on the original autoradiogram suggested the level of expression is more than an order (but less than two orders) of magnitude lower than from clone D8. The intensity of the lysozyme band from clone D8 is similar to that of the κ-chain band from X63.Ag8. The κ-chain

Figure 4.15. Comparative secretion-levels, of IgG from X63.Ag8, and of lysozyme from J558L transfectants containing two-gene, divergent, lysozyme expression-plasmids.

Cell lines, X63.Ag8, D8 and F3 (monoclonal, plys^{+}(d) transfectants of J558L), C6 (monoclonal, plys-SV40 transfectant of J558L) and untransfected-J558L, were brought to log-phase growth. Then, at 48-hour intervals, over a period of 5 days, aliquots of each cell suspension were removed, and proteins were biosynthetically-labelled by incubation with ^{35}S-methionine (see section 2.5.1.). Comparative levels of secreted proteins were determined, by immune-precipitation (see section 2.5.2.) from equal volumes of supernatants with the appropriate antibody, followed by 13% SDS-PAGE and autoradiography (see section 2.5.3.). IgG from the cell supernatants of X63.Ag8 was immune-precipitated with rabbit anti - mouse-IgG antibody, lysozyme from the cell supernatants of the transfectants was immune-precipitated with rabbit anti - chicken-lysozyme antibody and an 'immune-precipitation' was also performed on the supernatants from untransfected J558L, using rabbit anti - chicken-lysozyme antibody (negative control).

The photograph shows an autoradiogram of the SDS-PAGE gel. Tracks 1-5 contain immune-precipitates from radiolabelling incubations of day 1; tracks 6-10 contain immune-precipitates from radiolabelling incubations of day 3; tracks 11-14 contain immune-precipitates from radiolabelling incubations of day 5. Tracks 1, 6 and 11 contain X63.Ag8 immune-precipitates; tracks 2, 7 and 12 contain clone D8 immune-precipitates; tracks 3, 8 and 13 contain clone C6 immune-precipitates; tracks 4, 9 and 14 contain clone F3 immune-precipitates (although the lysozyme secreted is not visible on the photograph at the exposure shown) and tracks 5 and 10 contain J558L 'immune-precipitates'.

This series of experiments demonstrates that the radiolabelling / immune-precipitation method for relative quantitation of secreted protein can give reproducible results, even when performed at different times, providing the cells are carefully maintained. Also, it shows that clone D8 is the highest-level lysozyme-secreting transfectant, and this level of secretion appears to be at the same order of magnitude as Ig secretion from X63.Ag8 (discussed in section 4.7.6.).


1 2 3 4 5 6 7 8 9 10 11 12 13 14
66Kd→
29Kd→
14Kd→

contains four methionine residues (Svasti and Milstein, 1972) and lysozyme contains two methionine residues (Jung *et al.*, 1980). Therefore, in molar terms, this result suggests the level of lysozyme secretion is approximately twice that of the κ chain, or four times that of the complete $IgG_1(\kappa)$ molecule (γ_2,κ_2). Ig usually accounts for at least 10% of total cell-protein in myeloma cells and, therefore, this suggests that lysozyme is expressed at a high level. The level of expression is particularly high when considered in terms of a mammalian expression-system which has not undergone gene amplification.

In the above experiment, the assumption is made that the immune-precipitations are quantitative in the recovery of secreted proteins. To confirm this, crude cell-supernatants were run alongside immune-precipitates, obtained from the same volumes of cell supernatants, on a SDS-PAGE gel (see figure 4.16). Comparison of the signal intensities from the same protein bands show that the percentage recovery of IgG and lysozyme is similar, and near quantitative. Attempts to further increase the efficiency of recovery of the proteins, by altering the ratios and amounts of antibody and fixed *S. aureus* used in the immune-precipitations, were unsuccessful, suggesting the small losses are unavoidable transfer losses.

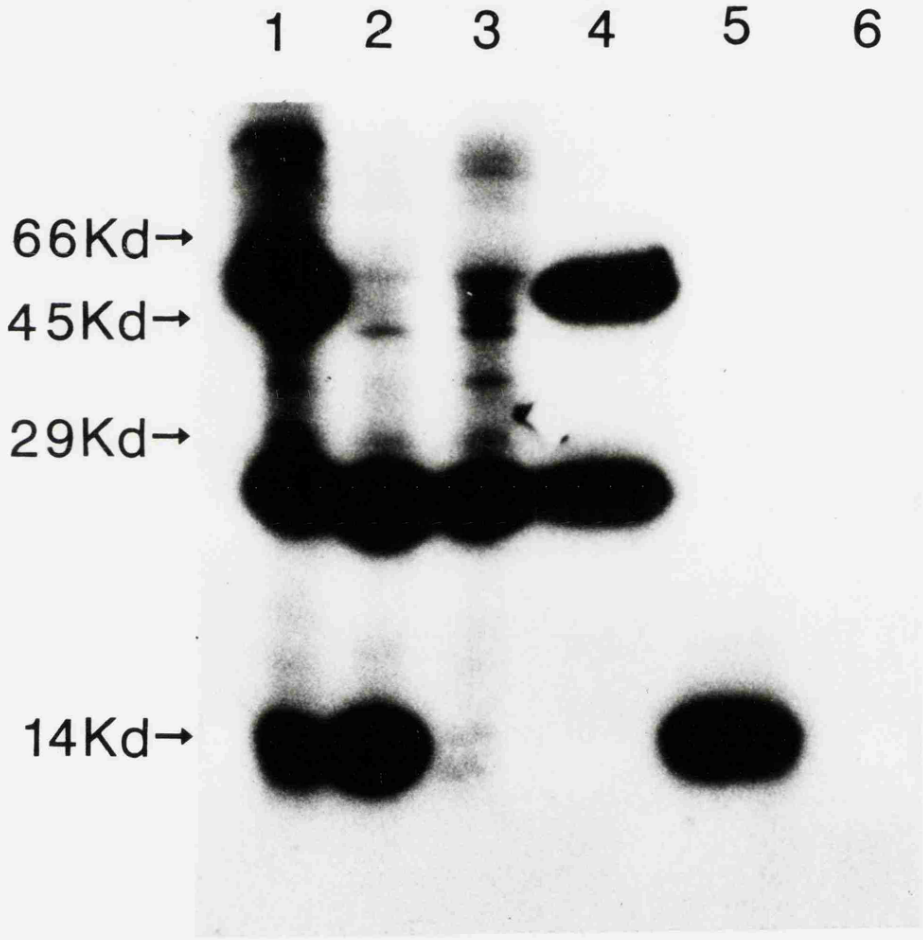
The crude supernatants from this experiment appear to be very clean. This is partially due to the high levels of lysozyme and Ig which make, in comparison, the secretion levels of other proteins appear small. Even so, if any heterologous protein can be produced and secreted at the same levels as lysozyme, its purification from the cell supernatant should be a simple task. This analysis also confirms that lysozyme is secreted, rather than its presence in the cell medium being due to cell lysis. Interestingly, in the X63.Ag8 crude-supernatant (track 1), there is a protein which comigrates with the lysozyme band in the J558L-transfectant tracks. This band (actually two bands) is also present in the untransfected-J558L crude supernatant (track 3) but at much lower levels. As discussed in section 4.7.4., these proteins might be truncated versions of Ig, which are expressed at different levels in different myeloma cell-lines. The autoradiogram also demonstrates the secretion of the endogenous, λ light-chain from J558L (track 3), in the absence of an IgH chain (see section 4.5.)

Figure 4.16. SDS-PAGE analysis illustrating near quantitative recovery of secreted proteins in immune precipitations.

Proteins of cell lines, X63.Ag8, D8 (monoclonal, plys^{+}(d) transfectant of J558L) and untransfected J558L, were biosynthetically-labelled for four hours with ^{35}S-methionine (see section 2.5.1.). Crude cell-supernatants, and immune-precipitates obtained from the same volumes of supernatants, were run on a 13% SDS-PAGE gel. The photograph shows the autoradiogram of this gel.

The tracks contain the following samples: track 1: X63.Ag8 crude-supernatant; track 2: clone D8 crude-supernatant; track 3: untransfected J558L crude-supernatant; track 4: IgG from X63.Ag8, immune-precipitated with rabbit anti - mouse-IgG antibody; track 5: lysozyme from clone D8, immune-precipitated with rabbit anti - chicken-lysozyme antibody; track 6: rabbit anti - chicken-lysozyme antibody 'immune-precipitate' from untransfected J558L (negative control).

This experiment shows that the recovery of secreted IgG and lysozyme, by immune-precipitation with rabbit anti - mouse-IgG antibody and rabbit anti - chicken-lysozyme antibody respectively, is almost quantitative.


1
2
3
4
5
6
66Kd→
45Kd→
29Kd→
14Kd→

The problem with measuring relative levels of expression, is that it can give misleading results. For example, the X63.Ag8 cell-line may have suffered a mutation, resulting in lower levels of IgG being secreted than was assumed. This would lead to an overestimation of the level of lysozyme secretion. Therefore, an attempt was made to quantify, absolutely, the level of lysozyme secretion, as described below.

Ideally, the amount of lysozyme accumulation in the medium under routine-culture conditions would be measured, using an ELISA or RIA. As an alternative, an assay system which did not require development, involving a biosynthetic-radiolabelling incubation in medium with a defined specific-activity of ^{35}S-methionine was used to estimate the level of lysozyme expression. Up to this point, radiolabelling incubations were carried out after addition of ^{35}S-methionine to methionine-free medium, which was supplemented with HS and FCS of undefined methionine-concentration. Hence, in these experiments the final specific-activity of the ^{35}S-methionine concentration was unknown and, therefore, it was only possible to determine comparative levels of secretion.

The above strategy necessitates the use of defined 'serum free' medium, which contains the growth factors required to allow the cells to grow in the absence of serum (alternatively, serum-supplemented medium that has been dialysed, to remove amino acids, could be used). It might be argued that serum-free medium is not ideal for obtaining efficient secretion of protein, because other members of the laboratory observed that J558L gradually perished when cultured in such medium. However, it is considered likely that the use of serum-free medium is unlikely to severely affect cells over the short time-period used for a radiolabelling incubation. Therefore, cell lines, D8 and X63.Ag8 were labelled in Iscove's defined medium (containing supplements; see section 2.5.1.). Aliquots of cell suspensions were removed at intervals over a 22-hour period, and the supernatants were separated. The proteins in the supernatants were immune-precipitated, and ^{35}S-methionine contents of the immune-precipitates were determined by scintillation counting. The counts were converted into weights of proteins (see table 4.2.), accounting for the proportion of ^{35}S-methionine as a fraction of the total methionine concentration (the specific-activity) in the medium used for radiolabelling, and also accounting for the methionine contents and molecular weights of the

proteins. Lysozyme contains two methionine residues (Jung *et al.*, 1980) and has a molecular weight of approximately 14KD. X63.Ag8 secretes an $IgG_1(x)$ protein called MOPC21; the complete molecule (γ_2, x_2) contains 26 methionine residues (9 in the γ chain [Adetugbo *et al.*, 1978] and 4 in the x chain [Svasti and Milstein, 1972]) and has an approximate molecular weight of 150KD.

Figure 4.17. shows the accumulation of immune-precipitable proteins in the media over the 22-hour incubation period. There was an initial lag-phase in the secretion of the proteins. This presumably reflects the time taken for the cells to recover from the washing steps used when setting up the incubation, and also the period taken for the radiolabel to be incorporated into protein, and for this protein to be secreted. The rates of protein secretion decreased between 12 and 22 hours incubation, which was probably due to exhaustion of the medium. Therefore, the rates of secretion were measured between the 8 and 12 hour time-points, during which the cells appear to have secreted protein under conditions nearest to optimum. Lysozyme secretion from cell line, D8, is estimated as $4\mu g/10^6$ cells/24 hours, and $IgG_1(x)$ secretion from X63.Ag8 is estimated as $6\mu g/10^6$ cells/24hours. These measurements are not precise, because the growth of cells during the radiolabelling incubation was not accounted for (values given per 10^6 cells at the beginning of the incubation), but this may be counterbalanced by the small transfer-losses which seem to be unavoidable in immune-precipitations (see above). However, I believe the quantitations are reasonable approximations, because these inaccuracies are likely to be quite small. Indeed, the the IgG secretion-level measurement obtained from X63.Ag8 is in the anticipated range. Hence, this suggests that the estimated level of lysozyme secretion is a good approximation. Also, the rational of estimating the level of μ-chain secretion from the pSV-Vμ1 transfectant, clone D4(A1), by comparison with secretion levels of IgG from X63.Ag8 (see section 4.6.), is validated.

The above result suggests the level of lysozyme secretion from clone D8, in molar terms, is about seven times greater than that of complete $IgG_1(x)$ molecules (γ_2, x_2), or 1.6 times greater than the 'average' (Mw = 37.5KD) Ig chain. This agrees with the result obtained by comparing band intensities on SDS-PAGE gels. The level, in terms of weight, of lysozyme secretion from clone D8, is slightly higher than the optimum

Table 4.2. Explanation of calculation of absolute levels of lysozyme secretion from clone D8 and IgG secretion from X63.Ag8.

Cells were plated out at 2 x 10^6 cells/ml in Iscove's medium (with supplements), containing ^{35}S-methionine (added from a stock at 50.4 TBq/mmol, 555Mbq/ml) at 3.7 Mbq/ml. As explained in section 4.7.6., 10µl aliquots of the cell culture were taken at time intervals and the counts of proteins ($IgG_1(x)$ or lysozyme) secreted from the cells were determined by immune-precipitation and scintillation counting, of the protein from the medium. The following list shows the factors that have to be taken into consideration when calculating the amount of lysozyme from a counts per minute (cpm) determination.

Consideration	Multiplication factor
a) Cpm → disintegrations per min. (dpm) (efficiency of scintillation counting = 79%)	x 1.23
b) Dpm → disintegrations per sec. (dps)	x 60
c) The number of ^{35}S-met. molecules (mmol) in the immune ppt. from a 10µl supernatant aliquot.	x (50 x 10^{-12})
d) The conc. of ^{35}S-met. in the medium (mmol/l)	x (1 x 10^5)
e) The conc. of total met. in the medium (mmol/l) accounting for specific activity ([met.] in Iscove's med. is 2x10^{-4} M and [^{35}S-met.] is 7.3x10^{-8} M)	x (2.7 x 10^3)
f) The conc. of lysozyme in the medium (mmol/l) (lysozyme contains 2 met. res.)	x 0.5
g) The amount of lysozyme in the medium (µg/ml) (lysozyme mol. wt. = 1.4 x 10^3)	x (1.4 x 10^3)
h) The amount of lysozyme produced in terms of cell density (µg/ml/10^6 cells) (conc. of cells : 2 x 10^6 cells/ml)	x 0.5

Hence, for lysozyme secretion, the cpm value is converted to µg/ml/10^6 cells by multiplying by 3.9. x 10^{-4}.

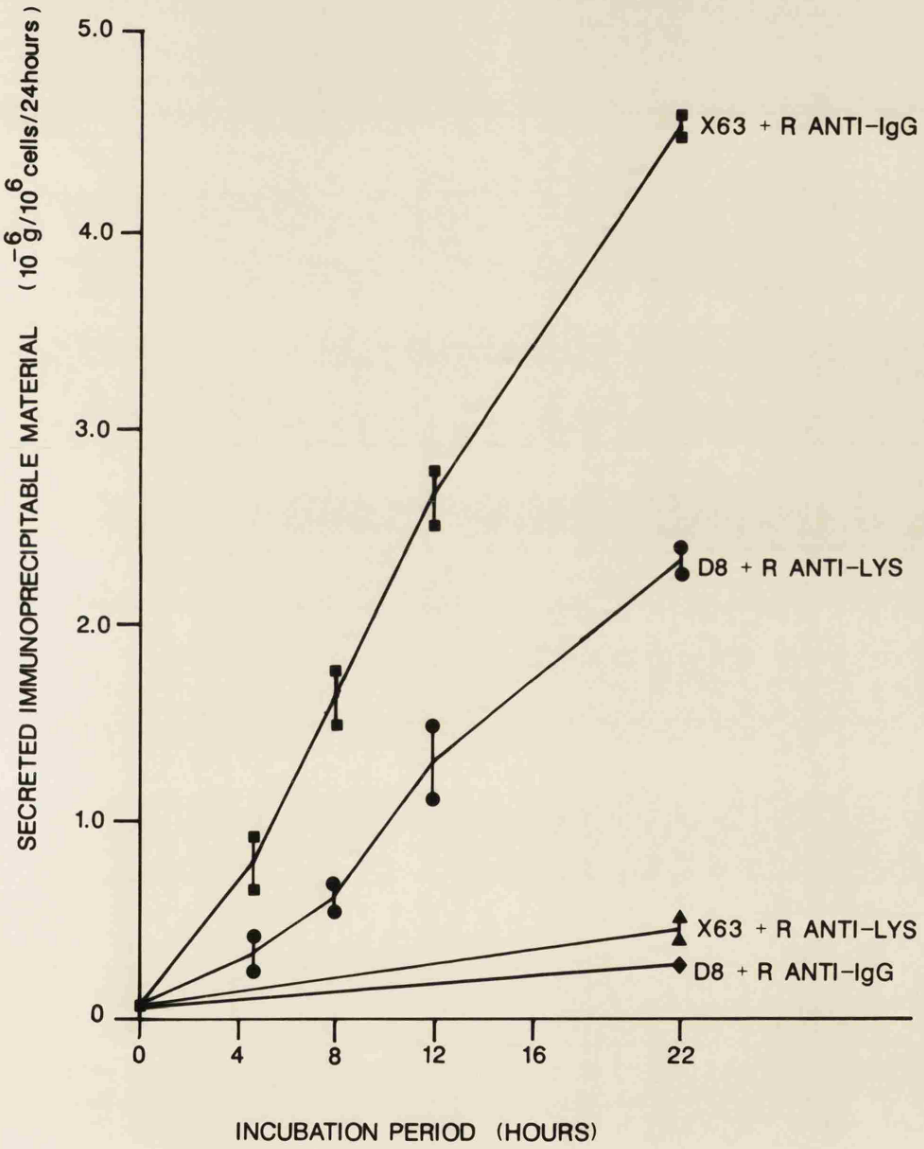
A similar calculation is used for determining the amount of $IgG_1(x)$ secretion from X63.Ag8. The only difference is that in step f) the multiplication factor is (1/24) because the IgG molecule, MOPC21 protein ($IgG_1(x)$), contains 26 methionine residues (see table 4.1.), and in step g) the multiplication factor is 1.5 x 10^4 because this is the approximate molecular weight of the IgG molecule. Hence, the cpm value is converted to µg/ml/10^6 cells by multiplying by 3.2 x 10^{-4}.

Abbreviations: min.: minute; sec.: second; met.: methionine; conc. concentration.

Figure 4.17. Time course for quantitatively-measured Ig and lysozyme secretion from J558L.

Proteins of cell lines, X63.Ag8 and D8 (monoclonal, plys^{+}(d) transfectant of J558L), were biosynthetically-labelled with ^{35}S-methionine in Iscove's defined medium (see section 2.5.1.). Aliquots of cell suspension were removed at time intervals, and the supernatants were separated from the cells. The amounts of IgG secreted from X63.Ag8 and lysozyme secreted from D8 at each time point were determined, by immune-precipitation from cell supernatants using rabbit anti - mouse-IgG and rabbit anti - chicken-lysozyme antibodies respectively, followed by scintillation counting of the immune-precipitates. For the final time-point, immune-precipitations were also performed using incompatible antibodies and supernatants (i.e. X63.Ag8 supernatant with anti-lysozyme and D8 supernatant with anti-IgG), to act as controls. Immune-precipitable counts were converted into weights of protein, taking into account the specific activity of radiolabelled methionine, and also the methionine content and molecular weights of the two proteins (see table 4.2.).

The graph shows the accumulation of secreted proteins in the medium with time. As described in section 4.7.6., an optimum rate of secretion from each cell-line was determined by measuring the level of secretion between the 8 and 12 hour time-points.


SECRETED IMMUNOPRECIPITABLE MATERIAL (10^{-6}g/10^6cells/24hours)
5.0
4.0
3.0
2.0
1.0
0
X63 + R ANTI-IgG
D8 + R ANTI-LYS
X63 + R ANTI-LYS
D8 + R ANTI-IgG
0
4
8
12
16
22
INCUBATION PERIOD (HOURS)

levels of secretion reported for some other heterologous proteins in myeloma cells. For example, Weidle and Buckel (1987) obtained expression of tissue-type plasminogen activator in J558L at levels up to 1μg/10^6 cells/24 hours. However, as lysozyme is a small protein, its expression level is even higher compared to that of such larger heterologous proteins, when expression levels are considered in molar terms. If the rate of transcription initiation is the limiting factor governing the level of expression from heterologous genes, then this suggests my expression plasmids are superior. However, the limitation on the secretion level of a heterologous protein is more likely to be at the post-transcriptional levels, particularly the post-translational level. Hence, it is perhaps unreasonable to expect the same molar-levels of secretion of larger, more complex proteins (see chapter 6).

As discussed in section 4.1., an assay for lysozyme activity exists which measures its ability to inhibit growth of the bacterium, *Micrococcus lysodeikticus*. Lysozyme hydrolyses the β(1→4) glycosidic-bonds between C-1 of N-acetylmuramate and C-4 of N-acetylglucosamine, which are the major polysaccharide components of the cell wall. The function of the cell wall is to confer mechanical support and, therefore, a bacterial cell devoid of its cell wall usually lyses because of the high osmotic-pressure inside the cell. As clone-D8 secretes recombinant lysozyme at high levels, it was expected that if the recombinant protein retains biological activity, it would be detected through the inhibition of *M. lysodeikticus* growth. Figure 4.18. shows that the supernatant of clone D8 does have antibacterial activity, presumably due to the presence of biologically-active lysozyme. As discussed in section 4.1., lysozyme is a simple protein which does not appear to undergo extensive post-translational modification. Even so, this result shows that myeloma cells do have the ability to produce recombinant proteins which possess biological activity. Hence, the protein most likely retains most of the features of the native protein.

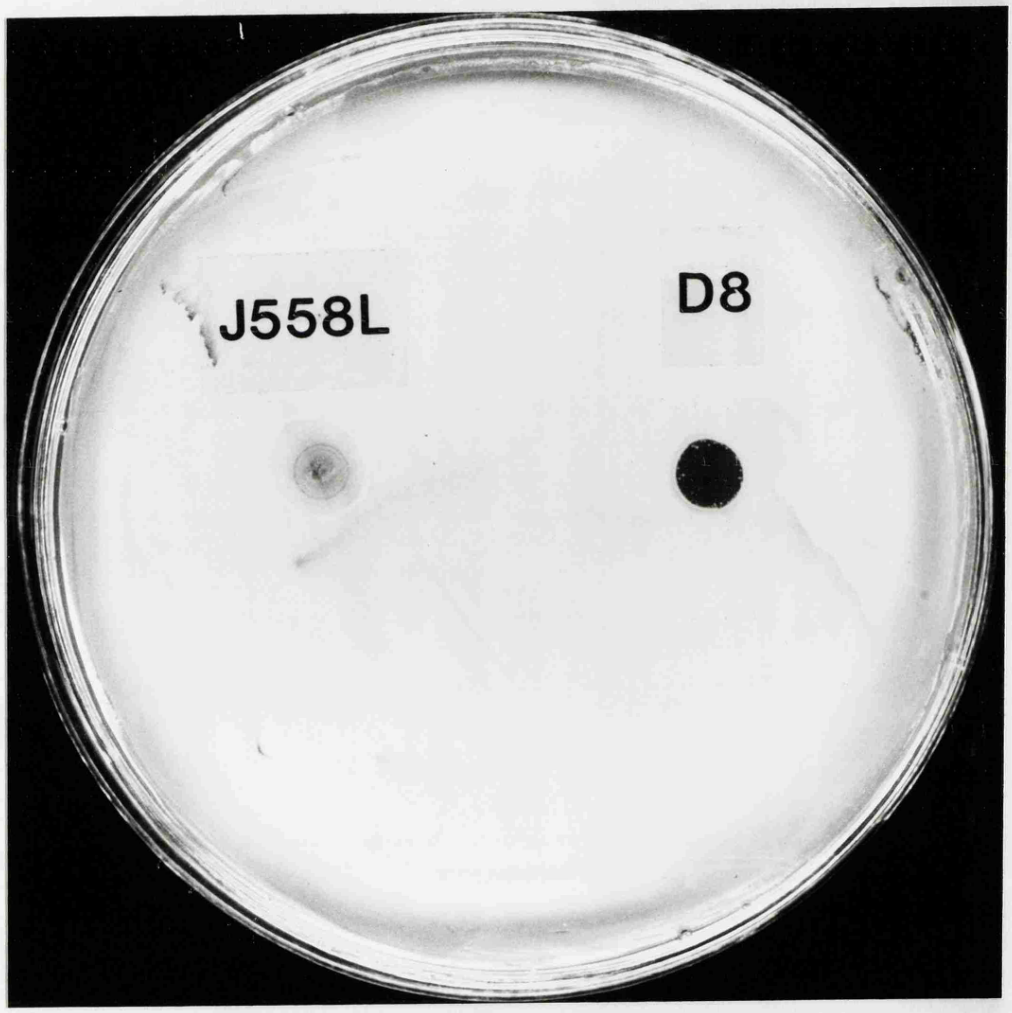
Antibacterial activity is a very rapid and simple assay for lysozyme activity. Therefore, providing the assay is sensitive enough to detect lower levels of lysozyme in cell supernatants, it could be used as a means of screening large numbers of clones to compare lysozyme secretion levels. The number of clones that can be screened using the biosynthetic-radiolabelling / immune-precipitation approach is limited because it is

Figure 4.18. Lysozyme produced in J558L has biological activity.

In view of the high level of lysozyme secreted from the the plys^{+}(d) transfectant of J558L, clone D8, it was considered that the level of lysozyme should be sufficient to give demonstrable anti-bacterial activity. Also, if a protein produced in a heterologous host can be shown to possess biological activity, it suggests that any essential post-translational modifications can be reproduced (although lysozyme does not, in fact, seem to be extensively modified).

An overnight culture of *Micrococcus lysodeikticus* (strain 1872) grown in LB-broth, was plated-out on a LB-agar plate. Once dried, 3µl aliquots of supernatants, from cultures containing clone D8 or untransfected J558L, at similar cell densities (≈ 10^{6} cells/ml), were spotted onto the agar plate. The plate shown was incubated at 30°C for 76 hours.

The photograph of the agar plate illustrates that a component of the supernatant from D8 (presumably lysozyme) prevents bacterial growth. In contrast, the bacterial lawn grows over the sample of applied supernatant obtained from untransfected J558L. There is, actually, a partial zone of clearing around the region where the untransfected-J558L supernatant was applied. This may be, in part, due to the penicillin and streptomycin in the cell-culture medium. However, the application of the supernatant sample may also disperse some of the the bacterial cells, as suggested by the greater thickness or 'crust' of the bacterial lawn at the edges of the applied supernatant spot.


J558L
D8

labour intensive. However, with this antibacterial-activity assay, more clones could be analysed and, therefore, the levels of expression obtained from plasmids containing different expression elements would have a greater statistical significance. Hence, it is likely to be useful for the further screening of expression elements in lysozyme expression-plasmids.

Clone-D8 was isolated towards the end of this project. The remaining parts of this chapter (sections 4.8. and 4.9.) describe attempts that were made to increase the level of lysozyme expression, using alternative expression-plasmids. Plasmids were constructed which differ in the expression elements flanking the lysozyme-cDNA fragment and/or the position and orientation of the lysozyme transcription-unit relative to the *gpt* transcription-unit. Although this work did not lead to any significant improvement in the average lysozyme expression-level, it illustrates some of the strategies which can be used when attempting to increase expression levels.

4.8. Tandem two-gene expression plasmids.

4.8.1. Introduction and Strategy.

It was decided to test the effect of altering the relative orientation of the non-selectable gene and the *gpt* gene, such that transcription of both genes was in the same direction. Plasmids containing this arrangement of transcription units are called tandem(t) expression-plasmids.

Three variant types of expression plasmid containing the tandem arrangement of transcription units were constructed. They differ in the positioning of the IgH-enhancer sequence(s) relative to the two transcription units. The plasmids were initially constructed as ovalbumin expression-plasmids (pEova(t), pEovaE(t) and povaE(t)). All three tandem, ovalbumin expression-plasmids contain a MT-I-promoter - ovalbumin cDNA - CMV-IE-poly(A) transcription-unit upstream of a ΔTK-promoter - *gpt* - SV40-downstream transcription-unit. However, because of the problems experienced in attempting to obtain ovalbumin expression from divergent expression-plasmids (see section 4.7.4.), they were then converted to lysozyme expression-plasmids for the purpose of assaying their efficiency. Also, as

Figure 4.19. Construction of pEova(t).

The strategy used to construct the tandem, ovalbumin expression-plasmid, pEova(t) (see section 4.8.1.), is illustrated and the details of this construction strategy are given below. pEova(t) contains the IgH enhancer directly upstream of an ovalbumin transcription-unit (utilising MT-I promoter and CMV-IE polyadenylation-signal expression elements), and a *gpt* transcription unit (identical to that in pΔTK-Egpt (see section 3.3.1.), except for restriction sites deleted at the junctions between the different fragments) directly downstream of the ovalbumin transcription-unit.

The pBR backbone sequences, which contain the ampicillin-resistance gene (amp), are represented by thin lines, the IgH gene major-intron fragments, carrying the enhancer (E), are represented by thin black boxes, the ΔTK promoter is represented by an unfilled box, the MT-I promoter is represented by a box filled with broken zig-zag lines, the CMV-IE promoter is represented by a box filled with closely-spaced, thin diagonal lines, the *gpt* sequence is represented by a box filled with thick diagonal lines, the ovalbumin cDNA is represented by a box filled with widely-spaced, thin diagonal lines, the SV40 sequences are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequence derived from the late region (SV40(L))) and the CMV-IE polyadenylation-signal is represented by a box filled with horizontal lines.

The plasmids which are illustrated completely are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Nde* I: N, *Eco* RI: R, *Hind* III: H, *Bam* HI: B, *Cla* I: C, *Sst* I: St, *Xba* I: X, *Sal* I: S, *Pst* I: Ps. A polylinker sequence is indicated as PL. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.

As the first step in the construction, the *Eco* RI site at the 5'-end of the MT-I promoter in pMT-Egpt (section 3.3.1.) was deleted, by cutting with *Eco* RI, filling-in with Klenow and then religating. Then the *Nde* I - *Bam* HI backbone was replaced by the *Nde* I - *Bam* HI backbone from pBR322-Δ*Hind* III (pBR322 from which the *Hind* III site was deleted, by *Hind* III digestion, filling-in and then religating). The resultant plasmid was digested with *Cla* I (and filled-in) and *Hind* III to remove the *gpt* sequence, downstream SV40 sequences and 0.4Kb of pBR sequence, and replaced with the 0.3Kb *Hind* III - *Sst* I (cut-back) CMV-IE-polyadenylation-signal - containing fragment from plys$^+$(d). Then a 1.4Kb *Hind* III ovalbumin-cDNA fragment (from pTK$_{\geq}$ov) was cloned into the *Hind* III site of the polylinker sequence separating the MT-I promoter and CMV-IE polyadenylation-signal, and a recombinant with the ovalbumin-cDNA insert in the sense orientation was identified. Then the *Eco* RI - *Nde* I backbone fragment from a Δ*Bam* HI, Δ*Hind* III derivative of pΔTK-Egpt (see section 4.7.1.), which contains *gpt*, SV40 (small t-antigen intron and early polyadenylation-signal) and part of the pBR backbone, was inserted between the *Nde* I and *Eco* RI sites, to produce pEova(t).

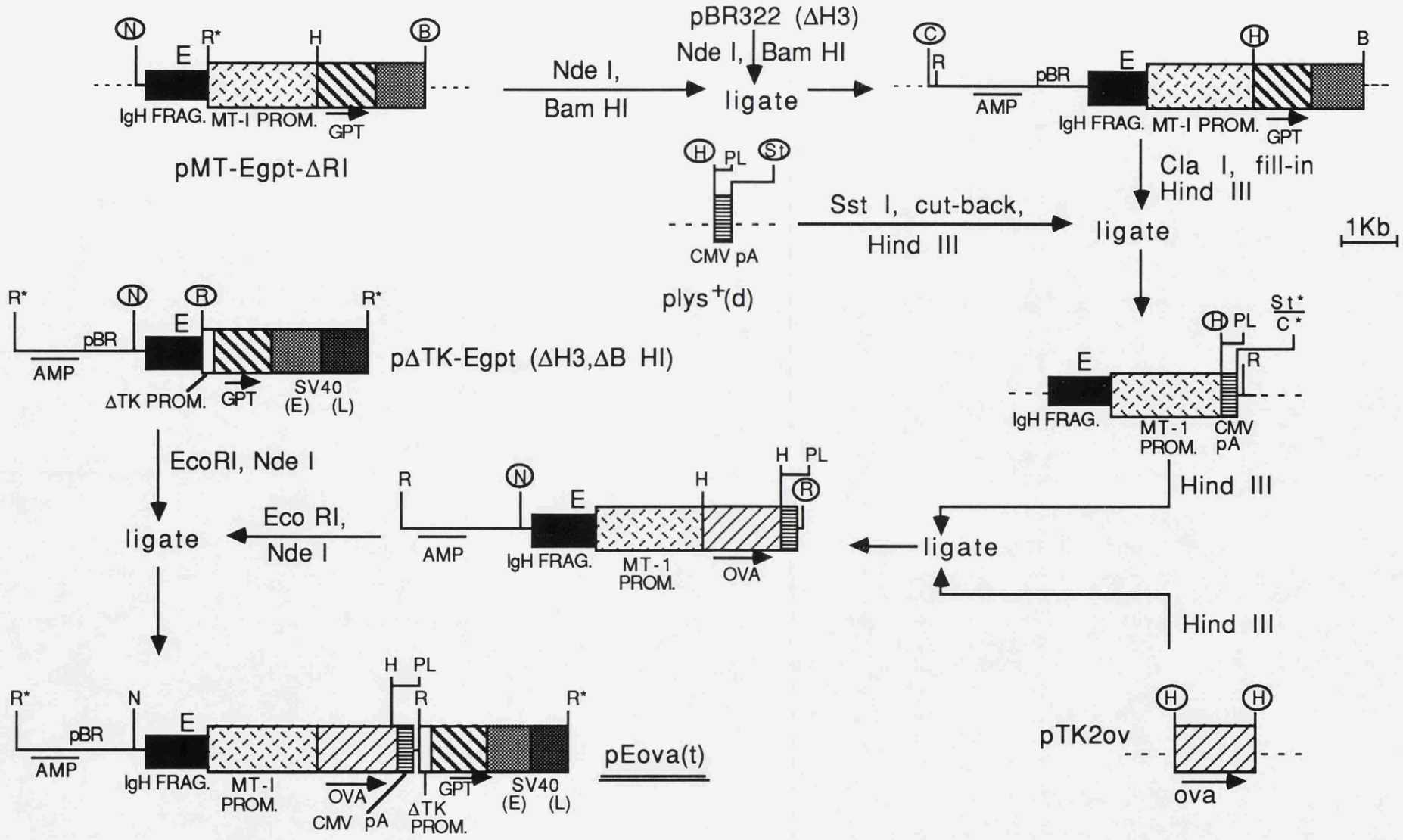

pMT-Egpt-ΔRI
IgH FRAG.
MT-I PROM.
GPT
Nde I,
Bam HI
pBR322 (ΔH3)
Nde I, Bam HI
ligate
AMP
pBR
Cla I, fill-in
Hind III
plys+(d)
CMV pA
Sst I, cut-back,
Hind III
1Kb
pΔTK-Egpt (ΔH3,ΔB HI)
ΔTK PROM.
SV40
(E) (L)
EcoRI, Nde I
Eco RI,
Nde I
OVA
pTK2ov
ova
pEova(t)

Figure 4.20. Construction of tandem, two-gene lysozyme expression-plasmids.

The figure illustrates the strategy used to construct the lysozyme expression-plasmids, pElys(t), pElysE(t) and plysE(t), and the details of the construction strategy are described below. All three plasmids contain, in tandem, a lysozyme transcription-unit (utilising MMTV-LTR and CMV-IE polyadenylation-signal expression elements) and a downstream, *gpt* transcription-unit (utilising a ΔTK promoter and SV40 downstream sequences), but differ in the position of the IgH enhancer sequence(s) (see section 4.8.1.)

The pBR backbone, which contains the ampicillin-resistance gene (amp), is represented by a thin line, the IgH gene major-intron fragments, containing the enhancer (E), are represented by thin black boxes, the MMTV-LTR is represented by a box filled with triangles, the ΔTK promoter is represented by an unfilled box, the MT-I promoter is represented by a box filled with broken zig-zag lines. the *gpt* sequence is represented by a box filled with thick diagonal lines, the lysozyme cDNA is represented by a box filled with widely-spaced, thin diagonal lines (extending from top left to bottom right), the ovalbumin cDNA is represented by a box filled with widely-spaced, thin diagonal lines (extending from bottom left to top right), SV40 sequences are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequence derived from the late region (SV40(L))) and the CMV-IE polyadenylation-signal is represented by a box filled with horizontal lines.

The plasmids illustrated in full are shown as linearised at an *Eco* RI site (although this site is actually deleted), for consistency with previous figures. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Bam* HI: B, *Xba* I: X, *Sal* I: S, *Pst* I: Ps, *Hind* III: H, *Nde* I: N, *Eco* RI: R, *Bgl* II: Bg, *Acc* I: A. A polylinker sequence is indicated as PL. Only sites relevant to the manipulations performed are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.

Construction strategies.

The strategy used for the construction of pEova(t) was described in the figure legend of figure 4.19. A second IgH-enhancer sequence was then inserted between the the two transcription units of pEova(t); pEova(t) was opened with *Eco* RI (and filled-in) and *Bgl* II, and then a *Pvu* II - *Bgl* II fragment from the Δ*Bam* HI, Δ*Hind* III derivative of pΔTK-Egpt (see section 4.7.1.) was inserted, to produce pEovaE(t)

The plasmid, povaE(t), was then made. Firstly, the *Hind* III site downstream of the ovalbumin-cDNA in pEovaE(t) was deleted, by partial *Hind* III - digestion, filling-in with Klenow and then religating. Then the *Nde* I - *Hind* III fragment, containing the IgH enhancer and the MT-I promotor, was replaced with the *Nde* I - *Hind* III from pMT-gpt (see section 3.4.), which contained only the MT-I promoter.

From these ovalbumin expression-plasmids, the corresponding lysozyme expression-plasmids (but containing the MMTV-LTR sequence instead of the MT-I promoter fragment) were made.

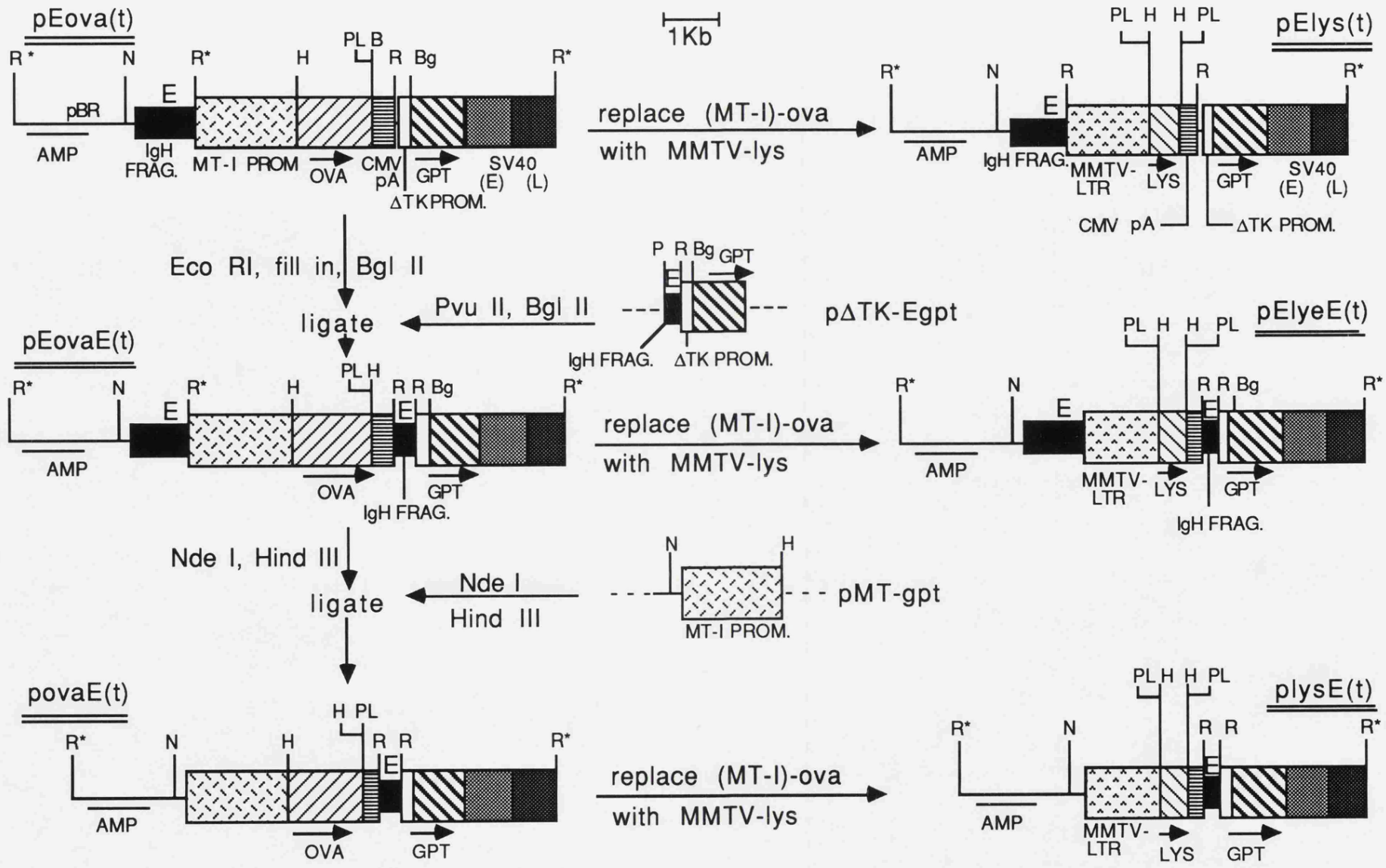
The ovalbumin-cDNA fragment in pEova(t) was removed by *Hind* III and *Bam* HI digestion, and then replaced with a *Hind* III - *Bam* HI lysozyme-

Figure 4.20. (cont.)

cDNA fragment taken from plys$^+$(d)-Δ*Hind* III (plys$^+$(d) from which the *Hind* III site at the 3'-end of the lysozyme cDNA was deleted, by partial *Hind* III - digestion, filling-in with Klenow and then religating). Then the MT-I promoter was removed by *Nde* I and *Hind* III digestion, and replaced by the *Nde* I - *Hind* III fragment from pMMTV-Egpt (see section 3.3.) which contains the MMTV-LTR, to generate pElys(t).

The plasmid, pElysE(t), was made from pEovaE(t), using exactly the same strategy as used for construction of pElys(t) from pEova(t).

The ovalbumin-cDNA fragment in povaE(t) was replaced by the lysozyme-cDNA as described for the above two constructs. Then the MT-I promoter was removed by opening with *Nde* I (and filling-in) and *Hind* III, and replaced with the *Eco* RI (filled-in) - *Hind* III fragment from pMMTV-Egpt (section 3.3.), which contained the MMTV-LTR. The resultant plasmid was called plysE(t).

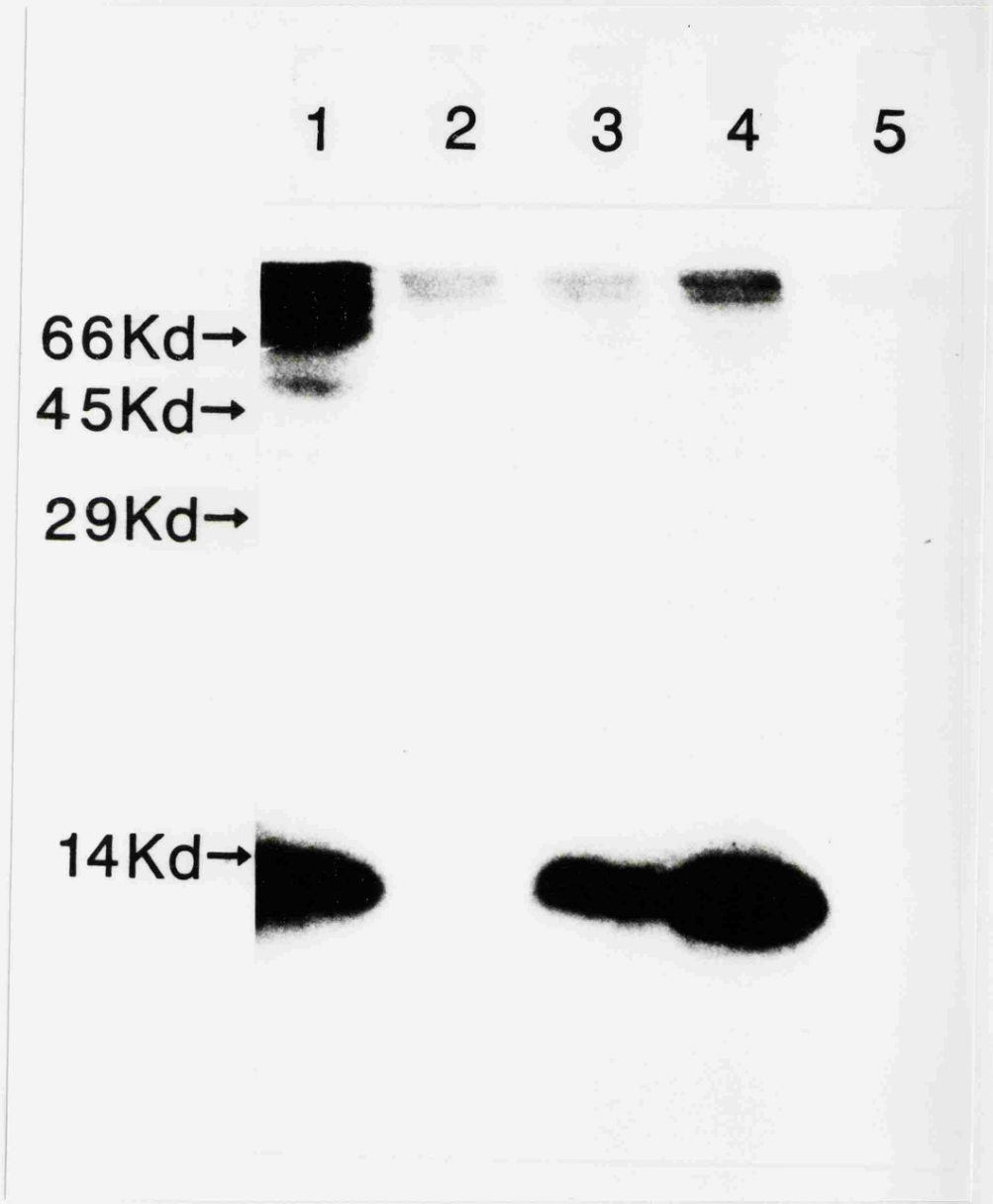

pEova(t)
R*
N
R*
H
PL B
R Bg
R*
pBR
E
AMP
IgH FRAG.
MT-I PROM
OVA
CMV pA
GPT
ΔTK PROM.
SV40 (E) (L)
1Kb
replace (MT-I)-ova with MMTV-lys
pElys(t)
PL H H PL
R*
N
R
R
R*
E
AMP
IgH FRAG.
MMTV-LTR
LYS
GPT
SV40 (E) (L)
CMV pA
ΔTK PROM.
Eco RI, fill in, Bgl II
ligate
Pvu II, Bgl II
P R Bg GPT
E
IgH FRAG.
ΔTK PROM.
pΔTK-Egpt
pEovaE(t)
PL H
R R Bg
OVA
GPT
IgH FRAG.
pElyeE(t)
Nde I, Hind III
Nde I
Hind III
N
H
MT-I PROM.
pMT-gpt
povaE(t)
H PL
R R
plysE(t)

the MMTV-LTR promoter element was considered to be the most active expression-element (see sections 3.3. and 3.4.), this was put upstream of the lysozyme-cDNA fragment. Hence, in each plasmid the MT-I-promoter - ovalbumin-cDNA sequence was replaced with a MMTV-LTR - lysozyme-cDNA sequence, resulting in the construction of pElys(t), pElysE(t) and plysE(t). As alluded to above with respect to the ovalbumin expression-plasmids, they differ in the position of the IgH-enhancer fragment(s) with respect to the two transcription-units. In pElys(t), the enhancer is positioned upstream of the ovalbumin transcription-unit; in plysE(t), the enhancer is sandwiched between the two transcription-units; and in pElysE(t), the enhancer is present in both of these positions. The strategies used in the construction of all these plasmids are illustrated in figures 4.19. and 4.20., and the construction strategies are described in detail in the accompanying figure-legends.

4.8.2. Lysozyme secretion from tandem two-gene expression-plasmids transfectants.

pElys(t), pElysE(t) and plysE(t) gave much higher stable *gpt*$^{+}$-transfection frequencies with J558L than the divergent two-gene lysozyme expression-plasmids (range: 10^{-4} - 10^{-5} transfectants per cell). Therefore, more than 200 colonies were pooled from each transfection, and the resultant populations were assumed to express statistically-representative levels of lysozyme. The lysozyme secretion-levels from the populations were compared, using the familiar procedure of biosynthetic-radiolabelling of protein, followed by immune-precipitation of lysozyme from cell supernatants and then analysis on SDS-PAGE gels. Figure 4.21 shows that the pElys(t) and plysE(t) transfectant populations expressed similar amounts of lysozyme. This seems to suggest that the IgH-enhancer works just as well when downstream of the transcription unit as when it is upstream. This might be expected, because the native position of the IgH enhancer in the IgH gene is within the major intron, which is 3' of the transcription-initiation site (see section 1.12.). It is surprising that the level of lysozyme expression from pElysE(t) transfectants was low in comparison.

Further screening of monoclonal transfectants (not shown) gave the same general results as with polyclonal populations; pElys(t) and


1
2
3
4
5
66Kd→
45Kd→
29Kd→
14Kd→

plysE(t) transfectant cell-lines were identified, which secreted levels of lysozyme similar to that from the plys$^+$(d) transfectant, clone F3 (see section 4.7.4.), which at that time (before isolation of the slow growing high-level-expressing clones described in section 4.7.6.) was the transfectant secreting the highest levels of lysozyme. Therefore, it appeared that there was no significant improvement in the levels of lysozyme expression from tandem expression-plasmids compared with divergent expression-plasmids, and consequently no further investigations were made on tandem expression-plasmids. It would now be interesting to see if further screening of pElys(t) or plysE(t) transfectants could lead to the isolation of clones secreting higher levels of lysozyme; particularly by screening colonies that grow slowly in XHMPA selective-medium, which was the avenue to isolation of the plys$^+$(d) transfectant, clone D8 (see section 4.7.6.).

4.9. Lysozyme expression-plasmids based around the IgH(μ) gene.

4.9.1. Introduction and strategy.

In section 4.6., it was shown that the pSV-Vμ1 transfectant of J558L, clone D4(A1), secretes the IgH(μ) chain, encoded by the transfected IgH(μ) gene, at levels similar to that of endogenous Ig-chains from the myeloma cell-line, X63.Ag8. Other pSV-Vμ1 - transfectant clones also secreted high levels of μ chain (see figure 4.5.). Myeloma cells are derived from plasma cells and, as discussed in sections 1.11. and 1.12., plasma cells are specifically adapted to the role of expressing and secreting large quantities of immunoglobulin. Therefore, it is no surprise that productive pSV-Vμ1 transfectants of J558L were easily isolated, considering the transfected IgH(μ) gene on this plasmid probably contains all or most of the native expression-signals.

It was decided to insert a lysozyme-cDNA into the transcribed region of a μ-gene, so that when this fusion gene (lys-μ) was introduced into J558L, it would allow generation of lys-μ transcripts. It was reasoned that sequences which are normally responsible for ensuring accumulation of high levels of Ig mRNA, might also contribute to high-level accumulation of mRNA from the fusion gene. If the lysozyme cDNA termination-codon was

preserved in the fusion transcript, such transcripts could encode lysozyme and, therefore, it was proposed that such a strategy might lead to high-level expression of lysozyme.

The only plasmid available containing a complete IgH-gene was pSV-Vμ1 (see section 4.3. and figure 4.4.). The most logical strategy was to insert a cDNA between the promoter and the V_H-region exon of the IgH gene. However, pSV-Vμ1 is a large plasmid containing multiple restriction-endonuclease recognition sites, and is not easy to manipulate in this way. In fact, during this study, Wiedle and Buckel (1987) reported the successful insertion of a cDNA sequence into a *Nco* I site (which contains the initiation codon) of the IgH(μ) gene in pSV-Vμ1, and were able to obtain transfectants expressing high levels of tissue-type plasminogen activator (as discussed in section 4.7.6.). This was a difficult manipulation involving linkers and, in addition, their reported construction strategy appears incomplete, because it does not describe how they overcame the presence of four *Nco* I sites in pSV-Vμ1 (presumably by partial digestion).

Using an alternative strategy, I made a two-gene expression plasmid (plysμ) which contains a lys-μ gene, in which most of the V_H-region and major intron of the μ-gene is deleted, and in which the IgH enhancer is relocated upstream of the IgH promoter. The construction strategy used in making plysμ is illustrated in figure 4.22., and the details of the construction strategy are given in the accompanying figure-legend.

The structure of the lys-μ gene in plysμ differs considerably from that of the complete μ-gene in pSV-Vμ1 (see figure 4.4.). In addition to the insertion of the lysozyme-cDNA fragment in place of (almost all) the V_H-region (L-VDJ$_2$), most of the Ig sequences extending to exon-1 of the Cμ-region are absent, including the enhancer, which is relocated upstream of the IgH promoter. It might be argued, therefore, that such alterations could affect the expression from the lys-μ gene. However, studies with the plasmid, pVμ1 (see figure 4.22.), seemed to suggest this was unlikely. pVμ1 is a μ-gene expression plasmid. It is, in effect, a derivative of pSV-Vμ1, in which, as in plysμ, the IgH enhancer is relocated upstream of the IgH promoter and in which most of the remaining major-intron sequences are deleted. The only difference between plysμ and pVμ1, as alluded to above, is the substitution of the lysozyme-cDNA fragment for (almost all) the V_H-

Figure 4.22. Construction of plysμ.

The figure illustrates the strategy used to construct plysμ. This plasmid contains a *gpt* transcription-unit, and also a second transcription-unit, which is convergent with the *gpt* transcription-unit at its 3'-end; this second transcription-unit contains fused lysozyme-cDNA fragment and IgH(μ)-gene sequences, and transcription is driven by an IgH enhancer and an IgH promoter positioned upstream. The rational behind the construction is outlined in section 4.9.1., and the details of the construction strategy are described below.

The pBR backbone, which contains the ampicillin-resistance gene (amp), is represented by a thin line, the IgH promoter is represented by a box filled with dots, the SV40 early-promoter fragment (also carrying the SV40 enhancer and late promoter) is represented by a box filled with squares, the *gpt* sequence is represented by a box filled with diagonal lines and SV40 downstream sequences are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early-polyadenylation signal (SV40(E)) is shaded less darkly than the sequence derived from the late region (SV40(L))). IgH-gene sequences other than the promoter, are shaded in black; intron sequences, including the relocated enhancer (E), are represented by the thinner boxes, and the exon sequences (the 3'-end of the V_H-region (drawn larger than to scale) which is fused with the 5'-end of the lysozyme-cDNA fragment, the constant regions exons (Cμ: 1, 2, 3 and 4) and membrane region exons (M)) are represented by the wider boxes.

The plasmids illustrated in full are shown as linearised at the site indicated. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Eco* RI: R, *Hind* III: H, *Sma* I: Sm, *Pvu* II: P, *Bam* HI: B, *Xba* I: X, *Mst* II: M, *Sal* I: S, *Nco* I: N, *Xho* I: Xh. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.

Construction strategy.

The strategy used to construct plysμ and the structure of the plasmids used in its construction are described below.

The first stage was to sequentially build up a sequence which contains, reading from 5'→3', the IgH enhancer, the IgH promoter, the lysozyme cDNA and a 5'-splice site, which could subsequently be fused with an intron sequence of the IgH-gene. Initially, a 0.5Kb lysozyme-cDNA *Hind* III fragment (from pTK_2lys) was inserted into the *Hind* III site downstream of the IgH promotor in pUC8-Ig (see section 3.3.1.). A recombinant containing the cDNA in the sense orientation with respect to the IgH promoter was identified. Then the *Hind* III site at the 5'-end of the lysozyme-cDNA insert fragment was deleted, by partial *Hind* III - digestion, filling-in with Klenow and then religating. Then the IgH-promoter - lysozyme-cDNA fragment was removed as an *Eco* RI (filled-in) - *Hind* III fragment and inserted into the *Bam* HI (filled-in) and *Hind* III sites of the pSV_2gptE' (see section 5.2.1. and figure 5.1.) backbone. This resulted in the formation of an IgH enhancer - IgH promoter - lysozyme-cDNA sequence. This was removed as an *Eco* RI - *Hind* III (filled-in) fragment and inserted

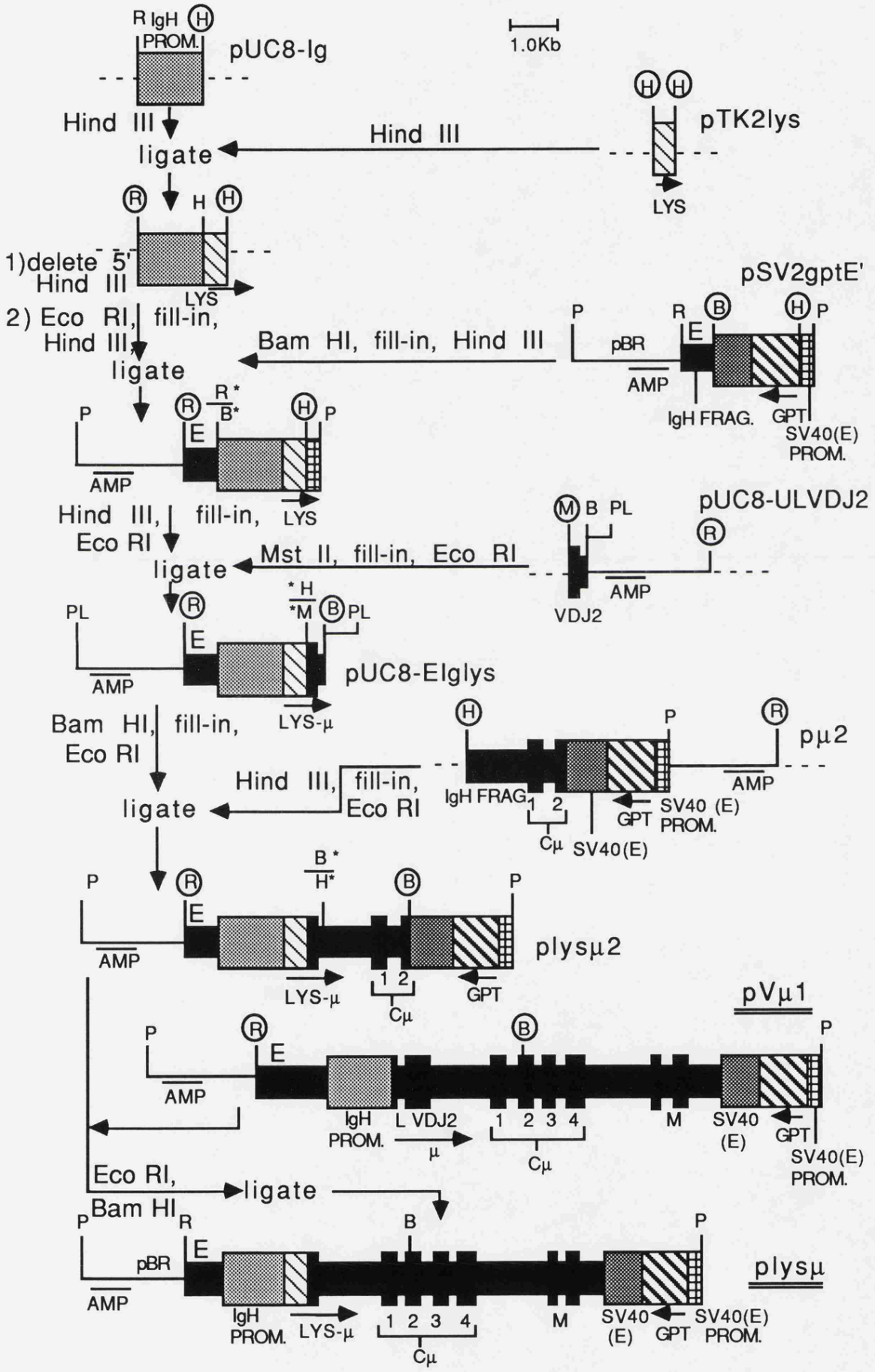
Figure 4.22. (cont.)
into the *Eco* RI - *Mst* II (filled-in) backbone of pUC8-LVDJ$_2$ (see figure 3.4.). This manipulation replaced the IgH promoter and almost all the V_H-region (LVDJ$_2$), and resulted in the fusion of the 3'-end of the lysozyme-cDNA fragment to the terminal four base-pairs at the 3'-end of the joining (J) segment of the V_H-region. This expected four base-pair fragment contains the sequence 5'-TCAG-3' (Neuberger, 1983), which is the core of the consensus sequence of a 5'(donor)-splice site (see section 1.9.3.). Hence, it should function as a splice site when introduced into J558L. The regeneration of the *Hind* III site from the fusion of filled-in *Hind* III and *Mst* II sites serves as an indication that the anticipated splice sequence is present in the resultant plasmid. This plasmid is called pUC8-EIglys.

The plasmid, pVμ1, was obtained from J.Ridge (lab. colleague). It contains, as in pSV-Vμ1 (see figure 4.4.), *a gpt* transcription-unit and an IgH(μ) transcription-unit, which are convergent at their 3'-ends. It differs from pSV-Vμ1, because IgH transcripts are not generated from a complete μ gene; the IgH enhancer and promoter sequences (the same fragment as in all other lys-μ constructs) are located upstream of the remaining μ-gene sequences. Also, some of the μ sequences are deleted. These are the J3 and J_4 exons (which are normally redundant in the μ-gene and are spliced out of transcripts as part of the major intron), and most of the major intron sequences (including the IgH enhancer) extending to the *Hind* III - site 5' of exon-1 of the Cμ-region.

The plasmid, pμ2, was obtained from S.Munson (lab. colleague). It was derived from pSV-Vμ1 (see figure 4.4.), by deletion of the 3.9Kb *Eco* RI fragment, followed by deletion of the 4.6Kb *Bam* HI fragment and then finally, by deletion of the *Hind* III site at the 3'-end of the SV40 early-promoter.

pUC8-EIglys (which is described above) was digested with *Bam* HI (and filled-in) and *Eco* RI to release the IgH enhancer - IgH promoter - lysozyme cDNA - splice-donor cassette from the polylinker sequence, and this was inserted into the *Hind* III (filled-in) - *Eco* RI backbone of pμ2, to produce plysμ2 . Hence, plysμ2 contains the same *gpt* transcription-unit as in pSV-Vμ1, and a second transcription-unit to drive transcription of a lysozyme-cDNA - partial-μ-gene fusion transcript. However, all μ-gene sequences downstream of the *Bam* HI site in exon-2 of the Cμ-region are absent in this plasmid.

plysμ, the plasmid containing complete (containing all 3'-μ - gene sequences) lys-μ sequences was then made. In effect, the 4.6Kb *Bam* HI fragment from pSV-Vμ1, containing all μ-gene sequences downstream of the *Bam* HI site in exon-2 of the Cμ-region, was inserted into the *Bam* HI site of plysμ2. This was actually performed by inserting the *Eco* RI - *Bam* HI fragment from plysμ2, which contains the lys-μ fusion gene, into the *Eco* RI - *Bam* HI backbone of pVμ1.


R IgH H
PROM.
pUC8-Ig
1.0Kb
Hind III
ligate
Hind III
pTK2lys
LYS
R H H
1)delete 5'
Hind III
LYS
2) Eco RI, fill-in,
Hind III,
ligate
Bam HI, fill-in, Hind III
pSV2gptE'
P
R B H P
E
pBR
AMP
IgH FRAG.
GPT
SV40(E)
PROM.
P
R R*
B*
H P
E
AMP
LYS
Hind III, fill-in,
Eco RI
ligate
Mst II, fill-in, Eco RI
M B PL
pUC8-ULVDJ2
R
AMP
VDJ2
PL
R
* H
*M B PL
E
AMP
pUC8-EIglys
LYS-μ
Bam HI, fill-in,
Eco RI
ligate
Hind III, fill-in,
Eco RI
H
P
R
pμ2
AMP
IgH FRAG
1 2
GPT
SV40 (E)
PROM.
Cμ
SV40(E)
P
R
B *
H*
B
P
E
AMP
plysμ2
LYS-μ
1 2
Cμ
GPT
pVμ1
P
R
B
P
E
AMP
IgH
PROM.
L VDJ2
μ
1 2 3 4
Cμ
M
SV40
(E)
GPT
SV40(E)
PROM.
Eco RI,
Bam HI
ligate
P
R
B
P
E
pBR
AMP
IgH
PROM.
LYS-μ
1 2 3 4
Cμ
M
SV40
(E)
GPT
SV40(E)
PROM.
plysμ

region of the reorganised μ-gene. Following transfection of pVμ1 into J558L, several transfectants were generated which expressed μ-chain at molar levels similar to the γ and κ chains from the myeloma cell-line, X63.Ag8 (analysis performed by S. Rutlidge, an undergraduate project student), which suggested the reorganisation of the μ-gene was not detrimental to expression. As it is the V_H-region which differs between different IgH genes, it is unlikely, when considered from an evolutionary viewpoint, to contain important expression-control elements. Therefore, the replacement of this sequence with a lysozyme-cDNA fragment was not expected to result in removal of expression-control elements. Hence, it might be expected, following transfection into J558L, that lys-μ transcripts could be efficiently generated from the lys-μ gene of plysμ, so as to give the potential for high-level lysozyme expression.

I then made several deletions in the μ-region of plysμ. Various deletions of Cμ-region exons were first made, and then further deletions were made in the downstream μ-region of the smallest exon-deletion derivative, plysμA (see figure 4.23.). It was envisaged that lysozyme expression from plysμ would be efficient and that one of the deletions might result in a significant decrease in lysozyme expression. If so, this might allow identification of a sequence, which normally contributes to a characteristic (e.g. highly-stable mRNA) that allows high-level IgH chain expression. Such a sequence could then be incorporated into other heterologous transcription-units of expression plasmids for use in J558L, and might lead to high-level expression of heterologous proteins on a more routine basis. Such a system would be better defined and easier to manipulate than using an entire Ig-gene.

4.9.2. Lysozyme secretion from lys-μ expression-plasmid transfectants.

plysμ and its μ-sequence deletion derivatives were transfected into J558L. The familiar method of biosynthetic-radiolabelling / immune-precipitation, followed by scintillation counting or SDS-PAGE, was used for comparative determinations of lysozyme secretion with that from the $plys^+$(d) transfectant, clone D8 (see section 4.7.6.).

Initially, analysis was performed on monoclonal transfectants and comparative lysozyme secretion-levels were determined by scintillation

Figure 4.23. μ sequence deletion derivatives of plysμ.

The figure illustrates the extent of the various deletions made in the μ-gene - derived sequences in plysμ. The rational behind making these derivatives are outlined in section 4.9, and the construction strategies are described in detail below.

The diagram of plysμ is shaded in the same way as described in figure 4.22. The restriction sites used for making the deletion derivatives of plysμ are illustrated. The position of the polyadenylation signals for the μ_m (pA_1) and μ_m (pA_2) forms of mRNA are also shown. The lines marked 'Δ' indicate the sizes of the μ-sequence deletions in each plasmid, and the restriction sites marking the boundaries of the deletions, whether reformed or not (indicated by an asterisk) during ligation, are shown. Restriction sites are abbreviated as follows: *Sma* I: Sm, *Hind* III: H, *Bam* HI: B, *Apa* I: A, *Kpn* I: K, *Xho* I: Xh.

Initially, several deletions were made in Cμ-region of plysμ as described below.

A deletion was made between the *Hind* III site at the 3'-end of the lysozyme-cDNA fragment and the *Apa* I site in exon-4 of the Cμ-region, in plysμ, to produce plysμA. This deletion could not be performed directly because the relevant restriction sites were not unique. Therefore use was made of pVμ1Δ2/4. (kindly donated by J. Ridge). pVμ1Δ2/4 is a derivative of pVμ1 (see figure 4.22.) in which sequences between the *Bam* HI and *Apa* I sites of the Cμ-region exons 2 and 4 respectively, are deleted. The fusion of the two sites regenerated a *Bam* HI site. Therefore, this allowed an *Eco* RI - *Hind* III (filled-in) fragment, from plysμ2, to be inserted into the *Eco* RI - *Bam* HI (filled-in) backbone of pVμ1Δ2/4, to produce plysμA.

The plasmid, plysμB, was constructed by cutting plysμ with *Bam* HI (and filling-in) and *Sma* I, and then religating the backbone. Hence, plysμB contains a deletion of all μ-sequences upstream of the *Bam HI* site inn exo-4 of the Cμ-region in plysμ.

The plasmid, plysμC, was constructed by, in effect, deleting between the *Hind* III site at the 3'-end of the lysozyme-cDNA fragment and the *Sma* I site of exon-1 of the Cμ-region in plysμ. This was actually performed by inserting the *Eco* RI - *Hind* III (filled-in) fragment from plysμ2 (which contained the IgH enhancer - IgH promoter - lysozyme-cDNA cassette) into the *Eco* RI - *Sma* I backbone of plysμ. Hence, plysμC contains part of exon-1 and complete μ-sequences downstream of it.

A further variation in the number of exons in the lys-μ transcript was made by, in effect, deleting between the *Bam* HI and *Apa* I sites of exons 2 and 4, and also by deleting between the *Hind* III site at the 3'-end of the lysozyme-cDNA fragment and the Sma I site in exon-1, of the Cμ-region in plysμ, to produce plysμD. This was actually performed by inserting an *Eco* RI - *Hind* III (filled-in) fragment from plysμ2, into the *Sma* I - *Eco* RI backbone of pVμ1Δ2/4. Hence, lys-μ fusion transcripts generated from plysμD are expected to contain part of exon-1 and a 2/4-hybrid exon.

Deletions of 3'-μ-gene sequences were then made in plysμA (see above), which is the smallest Cμ-region - deletion derivative of plysμ.

plysμA was cut with *Xba* I (and filled-in) and *Sma* I, and then the backbone was ligated. This deleted all downstream-μ sequences, except the μ_m polyadenylation-signal and sequences downstream. The resultant plasmid is called plysμA-Δ1.

plysμA was cut with *Xba* I and *Xho* I, the overhangs were filled-in

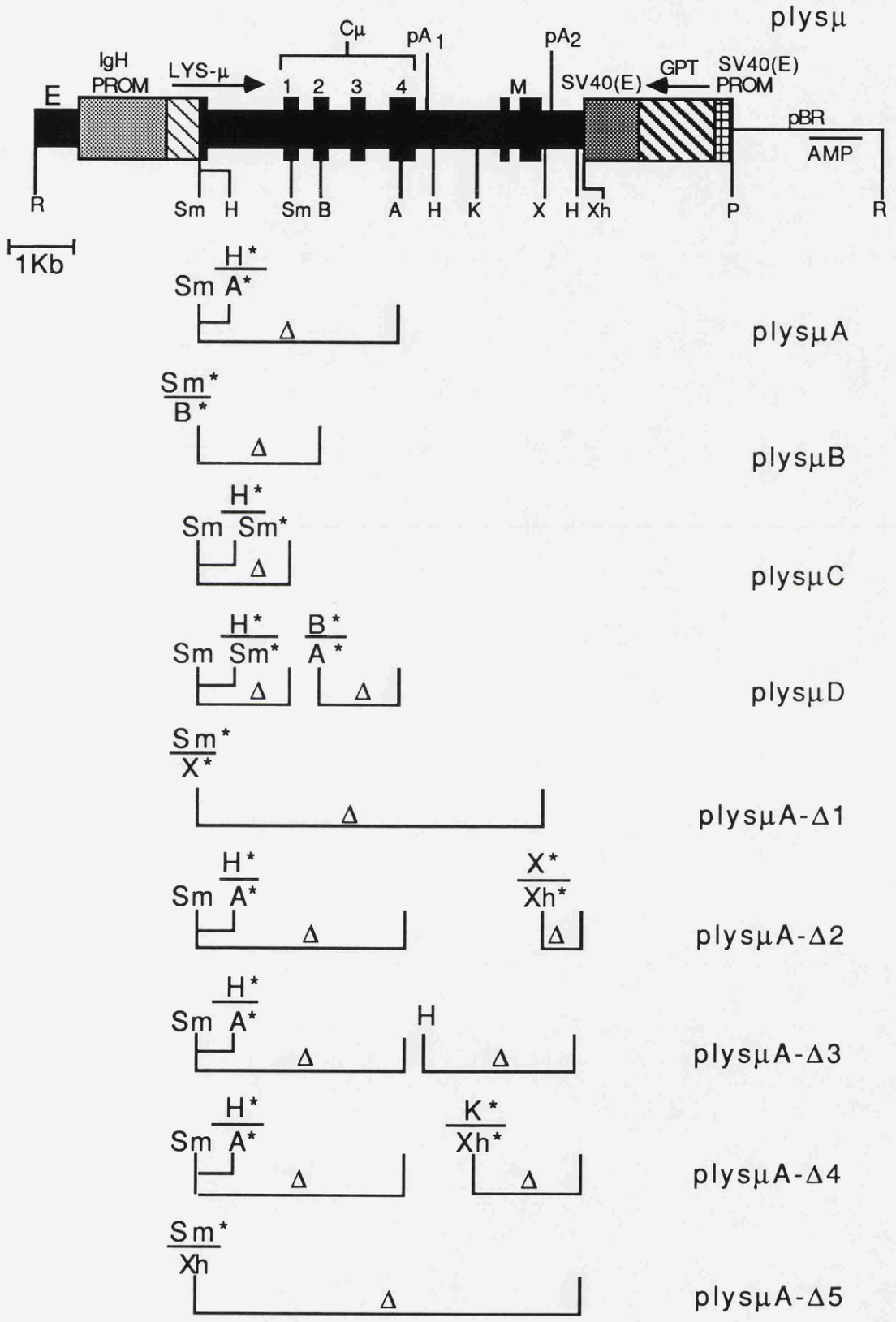
Figure 4.23. (cont.)

with Klenow and then the backbone was ligated. This deletion removed the μ_m polyadenylation-signal and μ-sequences downstream of it. The resultant plasmid is called plysμA-Δ2.

plysμA was partially digested with *Hind* III, filled-in with Klenow and then ligated. A recombinant was isolated, in which digestion had only occurred at the two *Hind* III sites in the μ sequences and not at a *Hind* III site at the 3'-end of the SV40 early-promoter. Therefore, this produced a deletion between the *Hind* III site downstream of the μ_m polyadenylation-signal and the *Hind* III site downstream of the μ_m polyadenylation-signal. This plasmid is called plysμA-Δ3.

plysμA was partially digested with *Kpn* I, and then completely with *Xho* I. Overhangs (both 5' and 3') were blunt-ended with Klenow, and then ligated. A recombinant was identified, in which a deletion had occurred between the *Kpn* I site, downstream of the μ_m polyadenylation-signal, and the *Xho* I site at the 3'-terminus of the μ sequences (i.e. the *Kpn* I site in the *gpt* sequence was not digested). This plasmid is called plysμA-Δ4.

plysμA was digested with *Xho* I (and filled-in) and *Sma* I, and then the backbone was ligated. The resultant recombinant, plysμA-Δ5, has no μ-sequences downstream of the lysozyme-cDNA fragment.


plysμ
Cμ
pA1
pA2
IgH
PROM
LYS-μ
1 2 3 4
M
SV40(E)
GPT
SV40(E)
PROM
E
pBR
AMP
R
Sm H
Sm B
A H K
X H Xh
P
R
1Kb
H*
Sm A*
Δ
plysμA
Sm*
B*
Δ
plysμB
H*
Sm Sm*
Δ
plysμC
H*
Sm Sm*
B*
A*
Δ
Δ
plysμD
Sm*
X*
Δ
plysμA-Δ1
H*
Sm A*
X*
Xh*
Δ
Δ
plysμA-Δ2
H*
Sm A*
H
Δ
Δ
plysμA-Δ3
H*
Sm A*
K*
Xh*
Δ
Δ
plysμA-Δ4
Sm*
Xh
Δ
plysμA-Δ5

counting of immune-precipitates. In view of the route by which the highest-level lysozyme-secreting cell lines were isolated (see section 4.7.6.), about half the colonies picked for expansion and analysis were those that grew slowly in XHMPA selective-medium during the selection process. Figure 4.24. shows, however, that the level of lysozyme secretion from all transfectants, and no matter what construct they contained, was disappointingly low. Ironically, the best expression, in the transfectants screened, was obtained from plysμA transfectants (and perhaps also its downstream-μ deletion derivatives); plysμA is the construct containing the least Cμ-region sequences of all the Cμ-region-exon deletion derivatives of plysμ.

At such low levels of secretion, it is difficult to differentiate, when using scintillation counting to quantify immune-precipitates, between actual secretion and variability of the background count. Hence, SDS-PAGE analysis of immune-precipitates was performed in order to confirm that lysozyme was being secreted. The stable *gpt*$^{+}$-transfection frequencies obtained with all the lys-μ constucts were high (see table 5.2). Hence, this allowed polyclonal populations, consisting of over 200 pooled transfectant-clones, to be used for the determination of, what were assumed to be, statistically-representative lysozyme secretion-levels. Also, SDS-PAGE analysis was performed on the immune-precipitates taken from several monoclonal plysμ-transfectants of J558L (different to those whose analysis is shown in figure 4.24.). Both these analyses (as shown in figure 4.25.) confirm that lysozyme can be expressed from the lys-μ genes, but support the conclusion that the general level of secretion is rather low, especially in comparison to the level of IgH expression that is obtainable from a transfected IgH-gene. The barely significant levels of lysozyme secretion detected in the analysis shown in figure 4.24., might have suggested that a mutation has occurred in the lysozyme coding-sequence, so as to completely obliterate expression from the lys-μ expression plasmids. The demonstration, in figure 4.25., that lysozyme is expressed, rules out this possibility. Autoradiogram A of figure 4.25. also suggests that plysμA is the most efficient lys-μ - gene - containing lysozyme expression-plasmid, which agrees with the analysis in figure 4.24. Autoradiogram B of figure 4.25. illustrates (as also seen in figure 4.5.) that expression is highly variable between different transfectant clones of

Figure 4.24. Comparative levels of lysozyme secretion from J558L monoclonal transfectants containing lys-μ expression-plasmids.

Proteins, of plysμ (and its μ-region - deletion derivatives) transfectants of J558L (see section 4.9.1.), were biosynthetically labelled with ^{35}S-methionine. Then the relative levels of lysozyme secretion were determined, by immune-precipitation, from equal volumes of cell supernatants, with rabbit anti - chicken-lysozyme antibody followed by scintillation counting.

The histogram illustrates the measured relative levels of lysozyme accumulation in cell supernatants. The groups show the levels from different monoclonal transfectants containing the same plasmids. The groups of transfectants contain the following plasmids: group 1: plysμ; group 2: plysμA; group 3: plysμB; group 4: plysμC; group 5: plysμD; group 6: plysμA-Δ1; group 7: plysμA-Δ2; group 8: plysμA-Δ3; group 9: plysμA-Δ4. The counts are also shown for an identical analyses on D8 (high level lysozyme-secreting plys^{+}(d) transfectant of J558L) and untransfected J558L, which act as positive and negative controls respectively.

The results from this analysis are discussed in section 4.9.2.

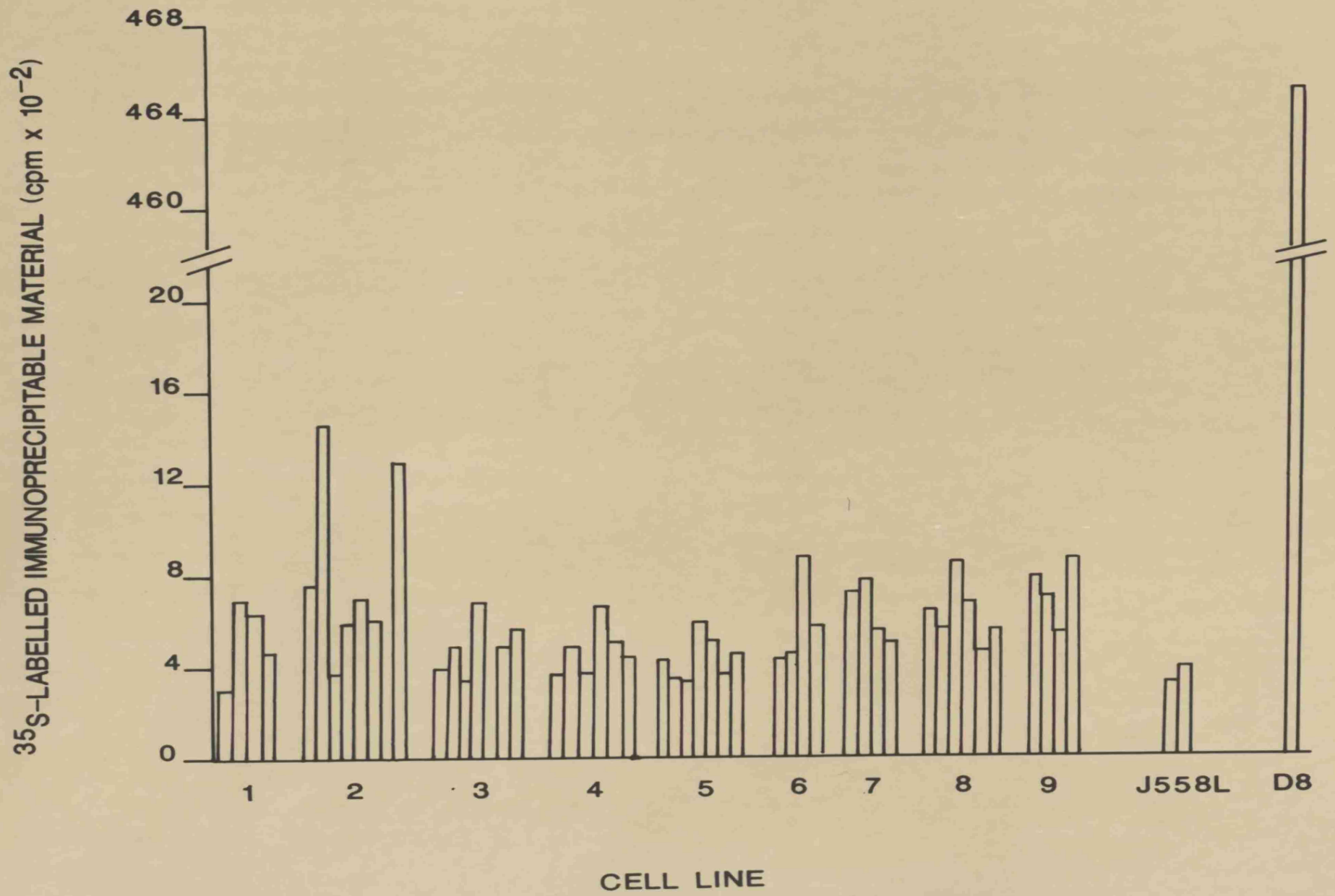
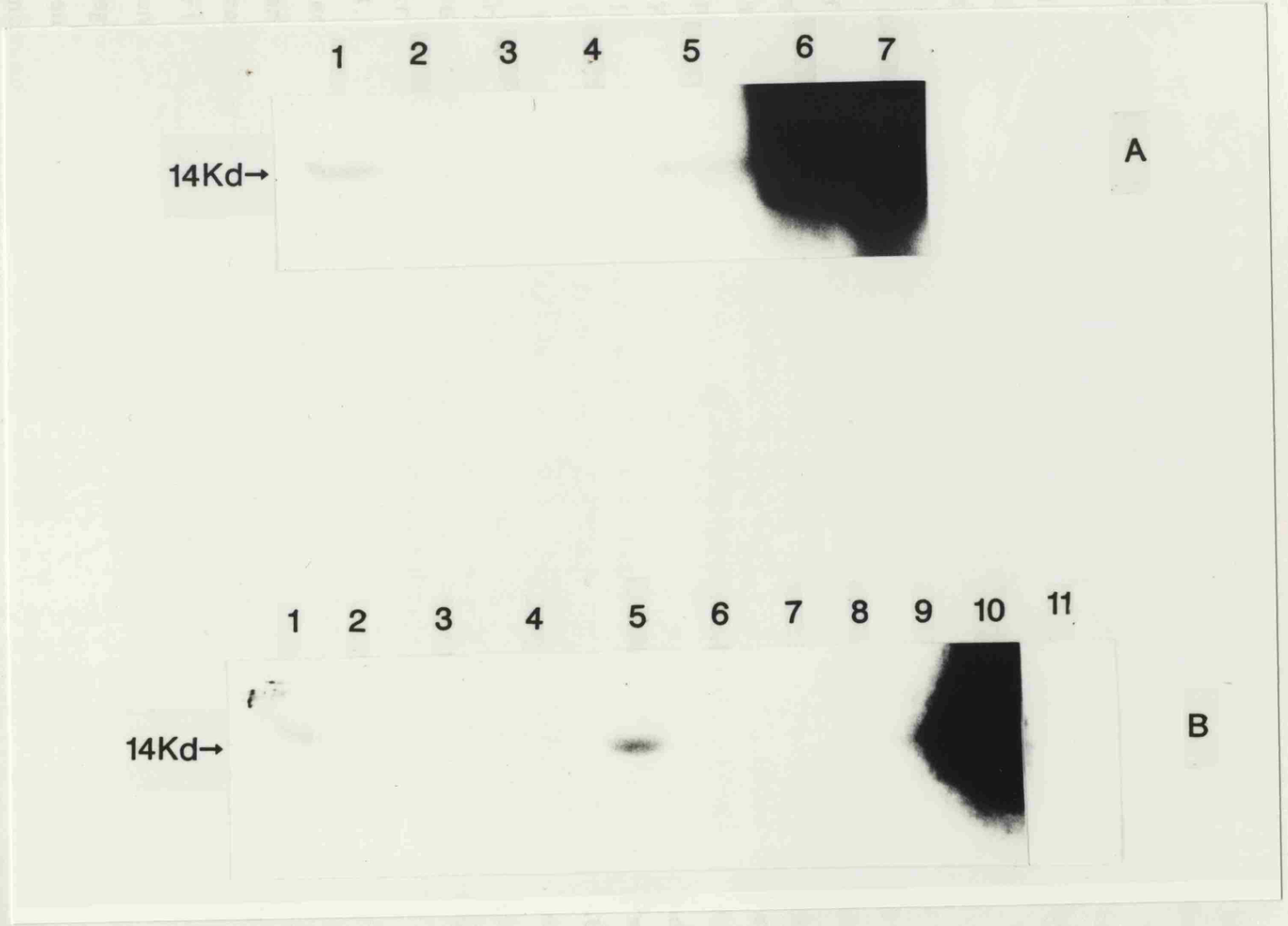

^{35}S-LABELLED IMMUNOPRECIPITABLE MATERIAL (cpm x 10^{-2})
468
464
460
20
16
12
8
4
0
1
2
3
4
5
6
7
8
9
J558L
D8
CELL LINE

Figure 4.25., Qualitative analysis of lysozyme secretion from J558L transfectants containing lys-μ expression-plasmids.

Figure 4.24. showed that the level of lysozyme secretion from plysμ (and its μ-region - deletion derivatives) transfectants of J558L is generally low. At such low levels of secretion, visual evidence is preferable to scintillation analysis for confirming that lysozyme secretion is, indeed, occurring. Hence, cells were biosynthetically-labelled with ^{35}S-methionine, and secreted lysozyme was immune-precipitated from equal volumes of cell supernatants with rabbit anti - chicken-lysozyme antibody. The immune-precipitates were run out on 13% SDS-PAGE gels. The photographs show the parts of autoradiograms of these gels, around the region of the expected 14KD lysozyme band.

Autoradiogram A shows a comparison of lysozyme secretion from polyclonal transfectant populations containing different lys-μ fusion gene expression-plasmids. The tracks contain lysozyme immune-precipitates from supernatants of J558L transfectants containing the following plasmids: track 1: plysμA; track 2: plysμB; track 3: plysμC; track 4: plysμD; track 5: plysμ; track 7: D8 (high level lysozyme-secreting plys$^+$(d) transfectant of J558L). Track 6 contains a X63.Ag8 crude supernatant, which is irrelevant to this analysis.

Autoradiogram B shows a comparison of lysozyme secretion from nine monoclonal plysμ-transfectants of J558L. The tracks contain lysozyme immune-precipitates from the following cell lines: tracks 1 - 9: plysμ transfectants; track 10: D8 (high level lysozyme-secreting plys$^+$(d) transfectant of J558L); track 11: untransfected J558L (negative control).


1 2 3 4 5 6 7
14Kd→
A
1 2 3 4 5 6 7 8 9 10 11
14Kd→
B

the same plasmid. Providing enough clones are screened, however, a plysμ-transfectant expressing easily-detectable amounts of lysozyme, although low by comparison to the plys⁺(d) transfectant, clone D8, can be identified.

As the lysozyme expression from all the constructs was so low, variations in expression against this low base are likely to be due to non-specific mechanisms, rather than the mechanisms which normally control the level of μ-chain expression. Hence, no information on sequences that normally ensure high-level μ-chain expression can be obtained from these results. Possible explanations for the low levels of expression are suggested below.

The lysozyme transcription unit of plysμA is quite similar to that of plys⁺(d) (which gave the highest-level lysozyme-expressing transfectant, clone D8). They both contain the IgH-enhancer and a promoter, upstream of the lysozyme-cDNA fragment, and a polyadenylation signal 'immediately' (≤ 0.5Kb) downstream of the cDNA. This may suggest that only a simple transcription-unit is required to obtain optimum expression. Also, the lysozyme-coding transcripts from these two plasmids are not expected to contain introns, which may indicate that expression of lysozyme is more efficient in the absence of introns. Alternatively, it may be an indication that the transcripts generated from the other lys-μ genes in the other constructs are inefficiently spliced; if so, they would be unlikely to be delivered to the cytoplasm because splicing of introns seems to be a pre-requirement for delivery of RNA to the cytoplasm (see sections 1.9.3. and 1.9.4.). Introns between exon-1 and exon-4 of the Cμ-region are likely to be recognised and spliced in J558L, because they contain the native μ-sequence splice sites. However, the intron between the lysozyme-cDNA fragment, and exon-1 of the Cμ-region might be less efficiently spliced. As described in the legend of figure 4.22., the 5'(donor)-splice site of this intron was made by fusing the terminal four base-pairs at the 3'-end of the V_H-region, to the 3'-end of the lysozyme-cDNA fragment. Although this sequence contains the three base 'core' of a consensus 5'-splice site, the absence of additional upstream-signals may have resulted in reduced efficiency of splicing. Also, there is no alternative 5'-splice site which could be used in this region, because the J_3 and J_4 exons, which are downstream of the V_H-region in the native IgH gene (which are not normally used, but might be used if there was no alternative splice-site) were

deleted during the construction of the lys-μ constructs. However, if inefficient splicing was the only reason why plysμA transfectants gave higher levels of lysozyme expression than transfectants containing other lys-μ expression-plasmids, higher levels of lysozyme expression should also have been obtained from plysμD transfectants, in which the intron between the lysozyme-cDNA fragment and exon-1 of the Cμ-region is also deleted.

The generally low level of lysozyme expression suggests there is some factor limiting the level of expression (relative to the level of μ-chain expression from a transfected μ-gene) from all constructs containing a lys-μ - gene. One possibility is that the incorporation of the lysozyme-cDNA fragment into a μ-gene transcript causes destabilisation. A similar observation was made by Hendricks and Banker (1987). They demonstrated that incorporation of the tissue-type-plasminogen-activator cDNA into the 5'-flanking region of an Ig gene destabilised the transcripts, such that the steady-state level of the hybrid message was only 1% of native Ig-mRNA. One possible explanation for the destabilisation, which is consistent with plysμA transfectants giving the highest level of expression, is that the hybrid transcripts are unstable because of long non-translated 3'-ends. In my lys-μ constructs, translation is expected to terminate at the termination codon at the 3'-end of the lysozyme cDNA. Hence, the entire μ-region of the transcripts is likely to be unprotected by ribosomes, perhaps leaving them susceptible to ribonuclease attack. If processing of the lys-μ transcripts generated from plysμA, occurs at the μ_m poly(A)-site, then the untranslated 3'-region would be, as alluded to above , only 0.5Kb in size (see Goldberg *et al.*, 1981). However, there may not be a complete correlation between the size of the non-translated μ sequence and transcript instability, because plysμ, which is expected to generate transcripts with the longest non-translated ends (3.3Kb; see Goldberg *et al.*, 1981), did not give the least expression amongst the polyclonal-transfectant populations tested (see figure 4.25., autoradiogram A).

In summary, it seems as though the approach of inserting a cDNA into an Ig gene, in order to obtain high-level expression from the cDNA, gives variable results. The results of Wiedle and Buckel (1987) demonstrate that this approach can be successful, but an empirical assessment of the level of expression fron a large number of cDNAs will be required to assess the versatility of this approach. However, it seems likely that only

through further understanding of the factors controlling Ig expression, and gene expression in general, will it be possible to introduce any coding sequence into an Ig gene, in the confident knowledge that it will lead to efficient expression following transfection

Although the level of lysozyme expression from plysμ and its deletion derivatives was low, the transfection experiments with this set of plasmids were also used to investigate the suspected phenomenon of 'transcriptional interference' in some of the two-gene expression plasmids, as discussed in chapter 5.

4.10. Summary.

Simple cotransfection was found not to be a satisfactory method for introducing a non-selectable gene into J558L; the cointegration, into the genome, of a non-selected gene with a selected gene (*gpt*) was not detected in the transfectants screened. In contrast, transfections with a two-gene (*gpt* and IgH(μ)) plasmid, pSV-Vμ1, resulted in high-level expression from the non-selected μ gene, in a satisfactory proportion of stable *gpt*$^+$-transfectants. This suggested that cointegration of a non-selectable gene with the *gpt* gene is much higher when the two genes are physically linked. Therefore, a series of two-gene plasmids were constructed. Expression plasmids for ovalbumin and lysozyme, in which the non-selectable gene encoding these proteins was divergent with a *gpt* gene, were initially constructed. Ovalbumin was not detected in the transfectants screened, but lysozyme was expressed and secreted from transfectants. Lysozyme expression-plasmids were then constructed in which the lysozyme and *gpt*-transcription units were arranged in tandem. Transfectants containing these tandem, lysozyme expression-plasmids seemed to express lysozyme as efficiently as those containing divergent, lysozyme expression-plasmids. However, more extensive screening was performed on transfectants containing the divergent, lysozyme expression-plasmids, and two clones expressing and secreting particularly high levels of lysozyme were isolated. One of these secretes lysozyme at higher levels (in molar terms) than the typical endogenous-Ig secretion level from myeloma cells, and the secreted lysozyme was shown to possess biological activity. Finally, attempts were made to express lysozyme from transfected plasmids in which

the lysozyme coding-sequence was inserted into an IgH(μ)-gene transcription unit, but this approach did not give high levels of lysozyme secretion in the transfectants tested.

CHAPTER 5: EXPRESSION INTERFERENCE IN TWO-GENE PLASMIDS.

5.1. Introduction.

Many of the expression plasmids used in this study contain two transcription-units; the first transcription-unit contains the selectable-marker gene (*gpt*), and the second transcription-unit contains the sequence encoding a protein of interest. In many two-gene expression vectors, it has been observed that adjacent transcription-units may interfere with each other, and result in reduced expression from one or both of the genes, compared with that from expression vectors containing the genes in isolation. This phenomenon is difficult to explain in mechanistic terms. However, three models, which are not necessarily mutually exclusive, have been proposed to explain it. One involves conformation, the second involves transcriptional interference, and the third involves competition for factors, as outlined below.

There are numerous lines of evidence supporting the view that the conformation of the chromatin within and flanking a gene may influence the potential for a gene to be expressed. As discussed in section 1.12. and demonstrated in chapter 3, the IgH enhancer is generally necessary for the production of a highly-active transcription unit, following transfection of a gene into a myeloma cell. However, the same element is dispensable from endogenous IgH-genes in myeloma cells without any effect on expression (see section 1.12.2.). This suggests that the enhancer must operate by setting up a stable transcription-enabling complex, which can be maintained in the absence of the enhancer. It is believed that the formation of such a complex must involve a change in the conformational arrangement of the chromatin, within and surrounding the gene. In addition, as described in section 1.9.1.3., the chromosomal location of a gene may affect its transcription rate, due to factors such as the methylation pattern and the non-random distribution of histone variants. Histones are generally regarded as structural or packing proteins, so if the variants differentially affect the transcription rate, it is likely to be through conformational changes in the chromatin. Also, Weintraub *et al.* (1985, 1986) emphasised the importance of DNA topology as an influence on the

transcriptional potential of DNA sequences. For example, they showed that transfected DNA is transcribed more efficiently when it is supercoiled rather than linearised. Despite evidence such as that described above, the mechanism by which active transcription-units become assembled, and also the three-dimensional arrangement of the chromatin and associated *trans*-acting factors in an active transcription-unit, are not well understood.

Although we are currently largely ignorant about variations in chromatin structure, there is much evidence that the chromatin structure associated with a transcribed gene is very different from that associated with non-transcribed regions of chromosomes. The chromatin structure associated with an active gene could conceivably extend into an adjacently-positioned transcription unit, and interfere with the assembly of an active transcription-enabling complex for that gene. For example, Emerman and Temin (1984, 1986) showed that in transfected Retroviral-vectors, which contained two transcription-units (one transcription-unit driven off the LTR, and the other off an internal promoter), when expression from one of the transcription units was selected for, there was normally suppression of expression from the the second transcription-unit. This effect was *cis*-acting (i.e. both genes had to be on the same vector), epigenetic and reversible. The effect was influenced by the promoter type but not by the relative strength of the promoters (the normal rate of transcription initiation from the promoter, in the transfected cell-type), nor, within limits, the distance between them. To explain this, they proposed that the establishment of one transcription-enabling complex changed the nearby chromatin-structure and precluded the efficient formation of an adjacent transcription-enabling complex. They suggested that the promoter specificity of the effect was due to different promoters binding different transcription-factors, which might distort the surrounding chromatin-structure to different extents. It was also noted that other sequences in the plasmid and surrounding genomic-sequences could influence the relative activity of the adjacent transcription-units. Therefore, it is hardly surprising that the phenomenon is difficult to understand. Briendl *et al.* (1984) and Hartung *et al.* (1986) also proposed similar explanations of transcriptionally-unfavourable topology, to account for how the insertion of a Retroviral vector nearby a collagen-gene promoter could greatly impair transcription from this promoter.

A second possible explanation for the inhibitory effect of active transcription-units on adjacent transcription-units, is one of transcriptional interference. This refers to the passage of a RNA-polymerase molecule through and beyond one transcription-unit, resulting in physical occlusion of the transcription of a neighbouring gene by a second RNA-polymerase molecule. This is a potential problem with most transfected genes, because eucaryotic expression-vectors do not normally contain recognised transcription-termination signals downstream of the polyadenylation signal; as discussed in section 1.9.2., few termination signals are well defined, but those which are defined are usually located far downstream of the polyadenylation signal. Therefore, they are unlikely to be contained in the generally small polyadenylation-signal-containing fragments used in expression vectors. Consequently, RNA-polymerase is likely to continue transcribing far downstream of the polyadenylation-site, until it falls off the template by chance, or meets some sequence with cryptic transcription-termination signals. However, particularly in expression vectors containing closely-positioned transcription units, there is the likelihood that RNA-polymerase could run into an adjacent transcription-unit before transcription termination occurs. Such an effect is direct, in contrast to the indirect conformation-based interference effect of adjacent, active transcription-units, as described above for the first model. The transcriptional-interference effect can easily be envisaged on circular vectors. It might also occur on integrated vectors, because they often integrate into the genome as concatamers containing multiple copies of vector (see section 1.6.); in addition to transcriptional interference between transcription units on the same plasmid, there could also be interference between transcription-units on adjacent plasmids of an integrated concatameric-unit.

Transcriptional interference might occur as one of two types. If the two transcription-units are arranged in tandem and in the same orientation, RNA-polymerase which initiates transcription at the upstream promoter, could continue transcribing into the promoter region of the downstream transcription-unit. Alternatively, if the two transcription units are convergent at their 3'-ends, RNA-polymerase transcribing in one direction might prevent the 'opposing' RNA-polymerase continuing far enough along the template to make a complete transcript.

The phenomenon in tandem transcription-units has been identified in both procaryotes (Hausler and Somerville, 1979; Adahya and Gottesman, 1982; Horowitz and Platt, 1982), in which it has been called promoter occlusion, and in eucaryotes (Cullen *et al.*, 1984; Proudfoot; 1986) in which it has been called transcriptional overlap interference. Cullen *et al.* (1984) showed that when there was no terminator sequence present between two transcription-initiation sites of tandem LTR-sequences in an Avian Retroviral-provirus, the transcription rate from the downstream (3'-) LTR was very low. However, when the upstream (5'-) LTR was inactivated, the 3'-LTR acted as an efficient promoter. They suggested that the RNA-polymerase which initiated transcription at the 5'-LTR promoter-element was obscuring the LTR promoter-element of the 3'-gene from independent transcription-initiation by other RNA-polymerase molecules. Proudfoot (1986) showed that by putting transcription-termination sequences (from mouse β-globin and sea-urchin H2A genes) between tandem α-globin - mini-genes, the transcriptional interference on the downstream gene could be relieved. These observations by Cullen *et al.* and Proudfoot led to the proposal that the absence or presence of active terminator-sequences can play an important role in controlling normal eucaryotic-gene expression.

I know of no published examples, in eucaryotes, of transcriptional interference between transcription units which are convergent at their 3'-ends. However, Ward and Murray (1979) demonstrated this form of transcriptional interference in procaryotes, between *trp* and λP_2 promoters in λ transducing-phage genomes. Gene expression was mutually impaired and, in fact, expression from the *trp* promoter was completely blocked. They also showed that the 'most effective' promoter (the one from which RNA-polymerase succeeded in completing the transcript following the assumed collision of polymerases) was not necessarily defined only by the frequency of transcription-initiation. This suggests that chromatin conformation must also influence the outcome in a situation involving transcriptional interference. Again, not enough detail is available on the topology to explain how opposing RNA-polymerase molecules can interfere with each other. The region containing RNA-polymerase, DNA, and nascent-RNA is called a 'transcriptional bubble' because it contains a locally melted 'bubble' of DNA, in which the DNA-helix is unwound. In some cases, it appears that this separation of the DNA-strands is not sufficient to allow

opposing polymerases on the different strands to pass.

The third explanation for reduced transcription from closely positioned promoters is one of competition. For example, the *trans*-acting factors (see section 1.9.1.), which interact with the *cis*-acting elements located in the promoter and/or RNA-polymerase, may be limiting in a localised region of the nucleus. If so, transcription initiation would be favoured from the promoter containing the *cis*-acting elements with the greatest affinity for such factors. However, there is little evidence to suggest this is the major explanation for the phenomenon.

In this chapter, I describe experimental results which suggested that some of the two-gene expression plasmids used in this study suffered from transcriptional interference, between transcription units at their 3'-ends. Attempts that were made to overcome this apparent problem are also described.

5.2. Circumstantial evidence for transcriptional interference.

5.2.1. Stable gpt+-transfection frequencies.

As discussed in section 3.1., the stable *gpt*$^+$-transfection frequency reflects the average level of XGPRT expression from each *gpt* expression-plasmid, which in turn probably reflects the efficiency with which complete *gpt*-transcripts are generated. The pPROM-Egpt constructs, which contain an IgH enhancer and a promoter, upstream of the *gpt* gene, gave the highest, stable *gpt*$^+$-transfection frequencies, generally greater than 10^{-4} transfectants per cell.

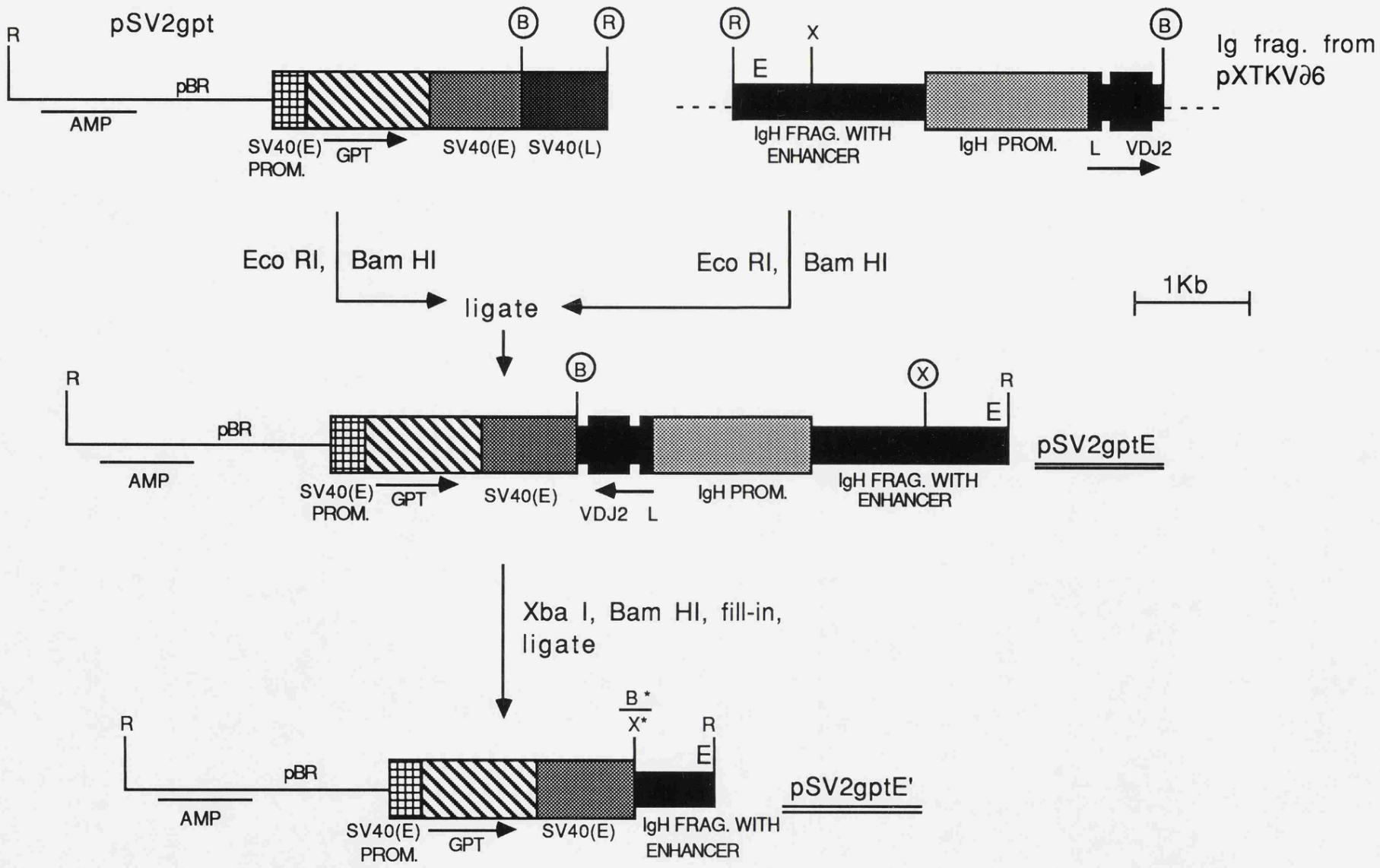
As the first step in the planned construction of a two-gene expression plasmid similar to pSV-Vμ1 (see figure 4.4.), which was to be used for obtaining expression from heterologous genes, the plasmid pSV$_2$gptE (see figure 5.1.) was made. Its construction involved the insertion of an *Eco* RI - *Bam* HI fragment (from pXTKVδ6), which contains an IgH enhancer - IgH promoter - L-VDJ$_2$ sequence, into the *Eco* RI - *Bam* HI backbone of pSV$_2$gpt. Surprisingly, when pSV$_2$gptE was transfected into J558L, it failed to give rise to stable *gpt*$^+$-transfectants in two spheroplast-fusion transfection experiments, whereas control-transfections performed with a pPROM-Egpt constuct, in parallel, gave typically high, stable *gpt*$^+$-

Figure 5.1. pSV_2gpt-derivative plasmids containing the IgH enhancer distant to the SV40 promoter.

The figure illustrates how the plasmids, pSV_2gptE and pSV_2gptE', were constructed (also see section 5.2.1.). A 3.65Kb fragment (containing the enhancer, the promoter and the V_H-region coding sequence of an IgH gene) was substituted for the SV40 late-region sequences in pSV_2gpt, generating pSV_2gptE. The IgH promoter was then deleted from pSV_2gptE, to produce pSV_2gptE', in which the IgH enhancer is located directly downstream of the *gpt* transcription-unit.

The SV40 early-promoter fragment (which also contains the SV40 enhancer and late promoter) is represented by a box filled with squares, the *gpt* sequence is represented by a box filled with diagonal lines, the SV40 fragments downstream of the *gpt* sequence are represented by darkly shaded boxes (the fragment containing the small t-antigen intron and the early polyadenylation-signal (SV40(E)) is shaded less darkly than the sequence derived from the late region (SV40(L))), the IgH promoter is represented by a box filled with dots, the potentially transcribed IgH-gene sequences are represented by black boxes (intron fragments, including that from the major intron, which contains the enhancer (E), are represented by the thinner black boxes, and V_H-region, leader (L) and (VDJ_2), exon sequences are represented by the wider black boxes) and the pBR backbone, which contains the ampicillin-resistance gene (amp), is represented by a thin line.

The plasmids are shown as linearised at an *Eco* RI site. Restriction sites are abbreviated as follows: *Pvu* II: P, *Hind* III: H, *Bam* HI: B, *Eco* RI: R, *Xba* I: X. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.


pSV2gpt
R
pBR
AMP
SV40(E) PROM.
GPT
SV40(E)
SV40(L)
B
R
R
E
X
B
Ig frag. from pXTKVδ6
IgH FRAG. WITH ENHANCER
IgH PROM.
L
VDJ2
Eco RI, Bam HI
Eco RI, Bam HI
ligate
1Kb
R
pBR
AMP
SV40(E) PROM.
GPT
SV40(E)
B
VDJ2
L
IgH PROM.
X
R
E
IgH FRAG. WITH ENHANCER
pSV2gptE
Xba I, Bam HI, fill-in, ligate
R
pBR
AMP
SV40(E) PROM.
GPT
SV40(E)
B *
X*
R
E
IgH FRAG. WITH ENHANCER
pSV2gptE'

transfection frequencies (see table 5.1.). In view of this result, another plasmid, pSV$_2$gptE' (see figure 5.1.), was then made, by deleting the IgH promoter and L-VDJ$_2$ - region from pSV$_2$gptE. This manipulation was performed by digesting with *Bam* HI and *Xba* I, filling-in with Klenow and then ligating the backbone. Hence, pSV$_2$gptE' contains a fragment carrying the IgH enhancer, directly downstream of the *gpt* transcription-unit. When pSV$_2$gptE' was introduced into J558L, it also did not produce any stable *gpt*$^+$-transfectants (see table 5.1).

One possible explanation for these results is that the IgH enhancer is not directly upstream of the SV40 early-promoter (which is expected to drive *gpt*-transcription). In contrast, the IgH enhancer is directly upstream of the SV40 early-promoter in the pPROM-Egpt construct, pSV-Egpt (see section 3.3.), which gave a high, stable *gpt*$^+$-transfection frequency. Therefore, the activating influence of the enhancer on the promoter might be severely reduced by the distance parting these transcription-regulating elements, resulting in there not being sufficient XGPRT-expression to generate *gpt*$^+$-transfectant colonies in XHMPA selective-medium. It is generally true that enhancers work more efficiently when directly adjacent to a promoter (Wasylyk *et al.*, 1984; Wasylyk and Wasylyk, 1986) and will preferentially activate the nearest promoter (de Villiers *et al.*, 1982; Wasylyk *et al.*, 1983; Kadesch and Berg, 1983). However, the IgH enhancer has been shown to be a particularly-strong activating element in lymphoid cells, and its action does not seem to be as adversely affected by distance from the promoter as for other enhancers, such as that of SV40. For example, Wang and Calame (1985) showed that the IgH enhancer can activate a V$_H$-region promoter separated by a distance of 17.5Kb. In this case, there was no inhibitory effect on the enhancer activation of the promoter when a second promoter was situated between the two elements. Moreover, the stable *gpt*$^+$-transfection frequency obtained with pSV-Vμ1 was only marginally lower than that obtained with a control, pPROM-Egpt-construct (see table 5.1.), and in pSV-Vμ1, the IgH enhancer is further removed from the SV40 early-promoter than in either pSV$_2$gptE or pSV$_2$gptE'. As can be seen in figure 4.4., in pSV-Vμ1, the IgH enhancer is located in the major intron of the μ gene and, therefore, is separated from the SV40 early-promoter by the pBR, IgH promoter and V$_H$-region sequences, upstream, and by 3'-Ig(μ) sequence, the downstream sequences of the *gpt*

Tables 5.1. and 5.2. Comparative stable-transfection frequencies of J558L to gpt⁺ following transfection with gpt expression-plasmids.

Tables 5.1. and 5.2. list stable *gpt*⁺-transfection frequencies obtained following transfection of J558L with various *gpt*-expression plasmids using spheroplast fusion. To ensure the selection conditions were comparable in all transfections, a parallel transfection was normally performed with a plasmid (control plasmid) capable of giving a high, stable *gpt*⁺-transfection frequency ($\approx 10^{-4}$ transfectants per cell). This plasmid is pΔTK-Egpt, unless otherwise indicated alongside the given transfection frequency. In most cases the pair of transfection experiments was repeated, and so the transfection frequencies from both sets of experiments are given When no control transfection was performed, this is indicated by N.D. (not done). When no stable transfectants were generated following a transfection experiment, this is indicated by N.T. (no transfectants). In situations when the transfection frequency was not recorded accurately, a + sign indicates that transfectant colonies were obtained.

Table 5.1. shows the stable *gpt*⁺-transfection frequencies obtained with a series of plasmids, which led to the following hypothesis; that expression from the *gpt* gene, in some of the two-gene expression plasmids containing transcription units which are convergent at their 3'-ends, can be inhibited by transcriptional interference.

Table 5.2., in contrast, lists the stable *gpt*⁺-transfection frequencies obtained with a series of plasmids, which are either contradictory to, or do not support, the hypothesis.

Table 5.1. Comparative transfection-frequencies to *gpt*$^+$ in J558L cells, which are consistent with the hypothesis that expression from *gpt* can be inhibited, by transcription units which are convergent with the *gpt* transcription unit at their 3'-ends, and which do not contain a transcription terminator.

Plasmid	Transfection Frequency (transfectants/cell x 10^4)	
	Test plasmid	Control plasmid
pSV-Vμ1	0.7	1.0
	0.5	4.0
pSV$_{\approx}$gpt	N.T. (< 0.001)	N.D.
	N.T. (< 0.001)	0.9
pSV$_{\approx}$gptE'	N.T. (< 0.001)	N.D.
	N.T. (< 0.001)	0.9
plys$^+$(d)	0.004	4.0
plys-SV40	0.05	4.0
plys-ISV40	0.004	4.0
plys-TK	0.02	4.0
plys-βglo	0.01	4.0
plysμ	0.5	1.0
plysμ2	N.T. (< 0.002)	4.0
plysμA	1.0	N.D.
plysμB	4.4	N.D.
plysμC	1.0	N.D.
plysμD	1.1	N.D.

transcription-unit and the *gpt* sequence, downstream. Hence, this suggests that the distance effect is of no more than minor importance.

pSV2gptE is a plasmid which is expected to drive transcription of a V_H-region (L-VDJ_2) sequence, because the IgH enhancer and the IgH promoter are directly upstream of this sequence. RNA-polymerase transcribing this sequence would be convergent with the 3'-end of the *gpt* transcription-unit (see figure 5.1). It now seems possible that transcripts could also be generated from the Ig fragment in pSV_2gptE', using cryptic transcription-initiation sites under the influence of the IgH enhancer, in a similar manner to that illustrated for p-Egpt (see section 3.6.). Such partial Ig-transcripts from both these constructs would have no available polyadenylation-signal. Hence, transcription termination might be particularly inefficient, because there is evidence that 3'-processing and transcription termination are, in some cases, coupled (see section 1.9.2.). It was proposed, therefore, that transcriptional interference (see section 5.1.), between opposing RNA-polymerase molecules on the *gpt* and Ig sequences inhibits completion of *gpt* transcription, resulting in the inability of the cell to produce sufficient XGPRT in order to generate selection-resistant colonies.

This hypothesis also seemed to be consistent with the observation that the divergent two-gene lysozyme expression-plasmids ($plys^+(d)$ and derivatives containing other polyadenylation-signals; see sections 4.7.1. and 4.7.5., and figure 4.14.) give low stable gpt^+-transfection frequencies (see table 5.1.). These plasmids contain divergent *gpt* and lysozyme transcription-units, positioned on either side of the IgH enhancer. Hence, RNA-polymerase molecules transcribing these transcription units might also be convergent at their 3'-ends. It was proposed that these constructs give measurable stable gpt^+-transfection frequencies whereas, pSV_2gptE and pSV_2gptE' do not, because they are more likely to generate complete *gpt*-transcripts. Firstly, the 3'-end of the lysozyme transcription-unit is further removed from the 3'-end of the *gpt* transcription-unit, due to the intervening pBR-derived sequences. Secondly, the plasmids contain a polyadenylation-signal downstream of the lysozyme coding-sequence, and as described in section 1.9.2., a functional polyadenylation-signal can form part of a transcription-termination signal. Hence, it is possible that transcription by some of the RNA-polymerase molecules producing lysozyme

transcripts can be terminated, before transcriptional interference occurs with the *gpt* transcription-unit.

As discussed above, another two-gene plasmid containing transcription units which abut at their 3'-ends, pSV-Vμ1, gives a high, stable *gpt*$^+$-transfection frequency (see table 5.1.). It is likely that Ig transcripts are produced as efficiently from pSV-Vμ1 transfectants as lysozyme transcripts are produced from divergent two-gene lysozyme expression-plasmids. The fact that high-level μ-chain expression was detected in several pSV-Vμ1-transfectants (see section 4.3 and figure 4.5.) suggests that the μ gene can be transcribed efficiently. Hence, if transcriptional interference is the explanation for the low, stable gpt$^+$-transfection frequencies produced from the above constructs, then this suggests that in pSV-Vμ1 transfectants, μ-gene transcription is efficiently terminated within the 3'-region of the μ gene, so as to avert transcriptional interference with RNA-polymerase on the *gpt* transcription-unit. The high, stable *gpt*$^+$-transfection frequency, from a plasmid containing a *gpt* transcription-unit which abuts at its 3'-end with a second transcription-unit containing a 3' - μ-gene sequence, does not only apply to pSV-Vμ1. High, stable *gpt*$^+$-transfection frequencies are also produced following transfection with plysμ and its Cμ-region deletion derivatives, plysμA, plysμB, plysμC and plysμD (see figure 4.23.). Interestingly, when the plasmid, plysμ2 (see figure 4.22.), was transfected into J558L, it did not give rise to stable *gpt*$^+$-transfectants (see table 5.1.). plysμ and plysμ2 are identical except that plysμ2 contains a deletion of all μ-gene sequence downstream of the *Bam* HI site in exon-2 of the Cμ-region. This result seemed to support the working hypothesis that sequences in the 3'-region of the μ gene can avert interference with *gpt* transcription. However, it must be conceded that the support of the hypothesis from the stable *gpt*$^+$-transfection frequencies obtained with plysμ and its derivatives, is not as strong as from the stable *gpt*$^+$-transfection frequency obtained with pSV-Vμ1. This is because the level of lysozyme expression from these constructs is low (see section 4.9.2.). It is suspected that this is due to instability of lys-μ transcripts. Without analysis of RNA levels, however, it can not be discounted that the rate of transcription of the lys-μ genes is low which, therefore, would not present the potential problem of transcriptional interference to the same degree.

As discussed in section 1.12.3., there is variability in the site of termination of IgH-gene transcription; in some myeloma cell-lines, termination occurs between the μ_m and μ_m polyadenylation-signals, whereas in the majority of cell lines, transcription terminates downstream of both polyadenylation-signals. However, whichever site of termination is favoured in J558L, it is possible that the above discussed plasmids, which contain the 3'-region of the μ-gene sequence, possess a termination signal. This is because, in addition to complete μ-gene sequences between μ_m and μ_m, they also contain μ-gene sequences which extend 0.7Kb downstream of the μ_m polyadenylation-signal.

5.2.2. RNA dot-blots.

In section 4.7.1. (also see figure 4.10.), I described the construction of the two-gene plasmid, plys$^-$(d). In this plasmid, the lysozyme coding-sequence is in the wrong orientation for lysozyme-coding transcripts to be generated through initiation at the the CMV-IE promoter. However, when RNA from plys$^-$(d) transfectants of J558L was hybridised to a lysozyme-RNA - complementary single-stranded DNA probe, it was seen that transcription of the lysozyme coding-sequence can occur (see figure 4.12.). Transcription could be initiated anywhere upstream or within the lysozyme-coding sequence, and without Northern-analysis and/or RNA-mapping to size and identify the span of the transcripts, it is not possible to be specific about their origin. However, one interesting possibility is that the transcripts are produced by continued transcription by RNA-polymerase II, after reading through the *gpt* transcription-unit. If the same transcripts are produced from plys$^+$(d) (an identical construct except the lysozyme coding-sequence has the opposite orientation; see section 4.7.1. and figure 4.10.), this could interfere with transcription along the lysozyme coding-strand and/or generate anti-sense transcripts to lysozyme-coding transcripts. Hence this phenomenon might limit the level of lysozyme expression from plys$^+$(d) transfectants. Overcoming such an inhibitory effect would be essential for obtaining optimum expression.

The kind of phenomenon described above does have precedence; Proudfoot (1986) presented evidence that transcripts can progress all the

way around an expression plasmid to interfere with a second transcription-unit.

5.3. Expression plasmids containing a transcription terminator.

If transcriptional interference was a problem in the divergent two-gene expression plasmids, it was expected that the introduction of an active transcription-terminator would avert the interference. The simplest assay available was to test if transcriptional interference on *gpt* transcription could be relieved, by analysing for an increase in the stable *gpt*$^+$-transfection frequency, following insertion of a terminator sequence in a position to prevent RNA-polymerase transcribing the non-selectable gene from progressing into the 3'-end of the *gpt* transcription-unit. If an effect was detectable, the transcription terminator could then be strategically inserted into expression plasmids, so as to minimise any transcriptional-interference which might be occurring on the transcription unit of the non-selectable gene, which encodes the protein of interest.

An available terminator-sequence was the relatively well-characterised sea-urchin histone (H2A) gene terminator-sequence [3'H2A] (Johnson *et al.*, 1986). This sequence was supplied on a 0.9Kb *Nru* I fragment (basically the *Taq* I fragment described by Johnson *et al.*, 1986). There are three sequence-motifs which make up the active terminator. One is found in the 3'-end of the H2A coding-region, and the other two are found in the 3'-flanking - region of the gene.

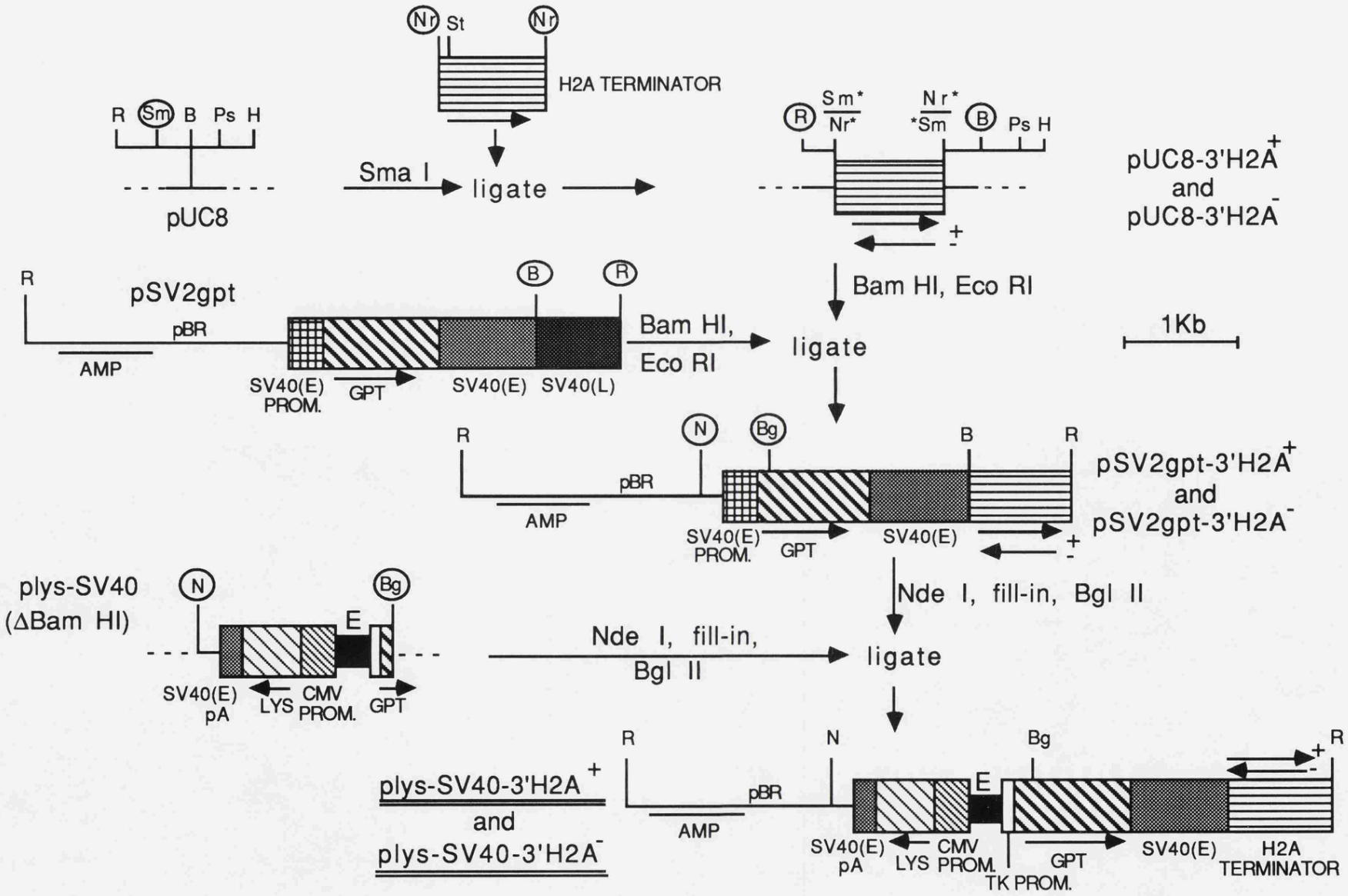
The above described *Nru* I fragment was, in effect, inserted, in both orientations, downstream of the *gpt* transcription-unit in the divergent two-gene lysozyme expression-plasmid, plys-SV40 (see figure 4.14.). This replaced the SV40 late-region sequence of plys-SV40 and resulted in production of plys-SV40-3'H2A$^+$ and plys-SV40-3'H2A$^-$ (see figure 5.2.). In fact, the *Nru* I fragment was first cloned into the *Sma* I site of the pUC8 polylinker-sequence and both orientations of the insert (pUC8-3'H2A$^+$ and pUC8-3'H2A$^-$) were isolated. These two orientations of the insert were then removed from the pUC polylinker as *Eco* RI - *Bam* HI fragments, and subcloned into the *Eco* RI - *Bam* HI backbone of pSV$_2$gpt, to generate pSV$_2$gpt-3'H2A$^+$ and pSV$_2$gpt-3'H2A$^-$ respectively. The *Bam* HI site downstream of the lysozyme-cDNA fragment in plys-SV40 was deleted, by partial *Bam* HI -

Figure 5.2. Construction of divergent, two-gene expression plasmids containing the sea-urchin H2A-gene transcription terminator.

The strategy used to, in effect, substitute the sea-urchin H2A gene transcription-termination sequence for the SV40 late-region sequence, in the divergent, two-gene expression plasmid, plys-SV40, is illustrated. This strategy is described in detail in section 5.3. Briefly, the following manipulations were performed. Firstly, the terminator sequence was cloned into the pUC8 polylinker-sequence, and both orientations of this insert were identified, making use of a *Sst* I site located close to the 5'-end of the insert. Then this fragment was substituted for the SV40 late-region sequence in pSV_2gpt, in both orientations. Finally, a second transcription unit, encoding lysozyme, was incorporated into the plasmids. This resulted in the generation of the constructs, plys-SV40-3'H2A$^+$ and plys-SV40-3'H2A$^-$, which differ only in the orientation of the terminator fragment.

The pBR backbone, which contains the ampicillin-resistance gene (amp), is represented by a thin line, the IgH gene major-intron fragment, containing the enhancer (E), is represented by a thin black box, the H2A terminator is represented by a box filled with horizontal lines, the fragment containing the SV40 early-promoter (and also the associated enhancer and late promoter) is represented by a box filled with squares, the ΔTK promoter is represented by an unfilled box, the *gpt* sequence is represented by a box filled with thick diagonal lines, SV40 fragments are represented by darkly shaded boxes (the fragments containing the small t-antigen intron, with (SV40(E)) or without (SV40(E)pA) the early polyadenylation-signal, are shaded less darkly than the fragment containing the late-region sequence (SV40(L))), the CMV-IE promoter is represented by a box filled with closely-spaced, thin diagonal lines and the lysozyme cDNA is represented by a box filled with widely-spaced, thin diagonal lines.

The plasmids illustrated completely are shown as linearised at an *Eco* RI site. Where a complete plasmid sequence is not illustrated, the non-illustrated sequence of such a plasmid is represented by a broken line. Restriction sites are abbreviated as follows: *Nru* I: Nr, *Sst* I: St, *Eco* RI: R, *Sma* I: Sm, *Bam* HI: B, *Pst* I: Ps, *Hind* III: H, *Nde* I: N, *Pvu* II: P, *Bgl* II: Bg. Sites used in each manipulation performed are circled, for emphasis. Only sites relevant to the manipulations are shown, for simplification. Sites deleted during construction of the plasmids are indicated with asterisks.


Nr St Nr
H2A TERMINATOR
R Sm B Ps H
pUC8
Sma I
ligate
R Sm* Nr* Nr* *Sm B Ps H
pUC8-3'H2A+
and
pUC8-3'H2A-
Bam HI, Eco RI
R
pSV2gpt
pBR
AMP
SV40(E) PROM.
GPT
SV40(E)
SV40(L)
B R
Bam HI, Eco RI
ligate
1Kb
R N Bg B R
pBR
AMP
SV40(E) PROM.
GPT
SV40(E)
pSV2gpt-3'H2A+
and
pSV2gpt-3'H2A-
Nde I, fill-in, Bgl II
plys-SV40
(ΔBam HI)
N Bg
E
SV40(E) pA
LYS
CMV PROM.
GPT
Nde I, fill-in,
Bgl II
ligate
plys-SV40-3'H2A+
and
plys-SV40-3'H2A-
R N Bg R
pBR
AMP
E
SV40(E) pA
LYS
CMV PROM.
TK PROM.
GPT
SV40(E)
H2A TERMINATOR

digestion, filling-in with Klenow and then religating. Then it was cut with *Nde* I (and filled-in) and *Bgl* II to release a fragment containing the lysozyme transcription-unit, the IgH enhancer and part of the *gpt* transcription-unit, and this was inserted into the *Nde* I (filled-in) - *Bgl* II fragments containing the pBR backbone and H2A-gene terminator from the two above pSV_2gpt derivatives, to produce the two constructs, plys-SV40-3'H2A$^+$ and plys-SV40-3'H2A$^-$.

If the terminator is orientation specific, then in plys-SV40-3'H2A$^+$ (in which the terminator is in the same orientation with respect to the *gpt* coding-sequence as it is to the coding sequence in the H2A gene), it was expected that *gpt* transcription would be efficiently terminated. However, this would not prevent transcriptional interference on the *gpt* transcription-unit by RNA-polymerase reading through the lysozyme transcription-unit, so little or no change in stable *gpt*$^+$-transfection frequency was expected compared with that obtained with plys-SV40. In contrast, in plys-SV40-3'H2A$^-$ transfectants, an active terminator was expected to terminate any extended 'lysozyme'-coding transcripts, resulting in a higher, stable *gpt*$^+$-transfection frequency. However, as shown in table 5.2., no stable *gpt*$^+$-transfectants were obtained following transfection with either of the terminator-containing constructs, plys-SV40-3'H2A$^+$ or plys-SV40-3'H2A$^-$, whereas the parent plasmid, plys-SV40, gave a low, stable *gpt*$^+$-transfection frequency typical of that usually observed with the divergent two-gene lysozyme expression-plasmids (see table 5.1.). Stable *gpt*$^+$-transfectants were also not generated, when the H2A-gene terminator sequence was inserted, in both orientations, at the same position as in the above constructs, in another divergent two-gene lysozyme expression-plasmid, plys-ISV40 (see section 4.7.5. and figure 4.14.).

It is possible that either the removal of the SV40 late-region fragment (which contains a polyadenylation signal) further reduces the potential for the *gpt* gene to be transcribed, and/or the terminator sequence is detrimental to *gpt* transcription. The results do not support the hypothesis that transcriptional interference is the major reason for low stable *gpt*$^+$-transfection frequencies. However, such an experiment can only be informative if a positive result is generated, because it is not known if the terminator is active in J558L. In the same way as transcription-initiation control is mediated through enhancer and promoter

elements, it is believed that control of downstream events associated with transcription occurs through the interaction of *trans*-acting factors with *cis*-acting sequence elements; if the appropriate combination of *trans*-acting factors required by the H2A-gene terminator is not present in J558L, the terminator sequence could not terminate transcription.

5.4. Deletions in the downstream region of the μ gene.

The data obtained from transfections with plasmids containing the H2A-gene terminator was ambiguous. Therefore, as an alternative strategy, an attempt was made to identify a transcription-termination sequence in the downstream region of the μ gene. As discussed in section 5.2.1., it was proposed that the high, stable *gpt*$^+$-transfection frequencies, obtained with two-gene plasmids containing the downstream region of the μ gene, is made possible by efficient termination of 3'-μ-sequence transcription which, therefore, prevents transcriptional interference with the *gpt* transcription-unit.

Deletions were made of various fragments from the downstream μ-region sequence in plysμA, to produce plysμA-Δ1, plysμA-Δ2, plysμA-Δ3, plysμA-Δ4 and plysμA-Δ5 (see figure 4.23.). plysμA contains a *gpt* transcription-unit, and a lys-μ transcription-unit which contains the 3'-region of the μ gene. These transcription units abut at their 3'-ends. If any of the downstream - μ-sequence deletions caused a dramatic reduction in the stable *gpt*$^+$-transfection frequency, this might have led to the identification of the sequences which were proposed to be preventing transcriptional-interference on the *gpt* transcription-unit (see section 5.2.1). However, table 5.2. shows that none of the deletion-derivatives gave a large reduction in stable *gpt*$^+$-transfection frequency. The most surprising result is the high, stable *gpt*$^+$-transfection frequency obtained with plysμA-Δ5. This plasmid contains a deletion of all the 3'-μ sequences (i.e. μ_m and μ_m polyadenylation-signals and the downstream region) and, therefore, the 3'-end of the lysozyme-cDNA fragment abuts directly with the 3'-end of the *gpt* transcription-unit. As with plysμ2 (see figure 4.22. and section 5.2.1.), it was expected that the deletion of complete 3-μ sequences would result in loss of the capability to generate stable *gpt*$^+$-transfectants. The fact that transfectants are generated at a high

Table 5.2. Comparative transfection-frequencies to *gpt*$^+$ in J558L cells which do not support the transcriptional-interference hypothesis.

Plasmid	Transfection frequency (transfectants / cell x 10^4)	
	Test plasmid	Control plasmid
plysμA-Δ1	0.3	1.0
	0.8	4.0
plysμA-Δ2	0.3	1.0
	0.3	4.0
plysμA-Δ3	0.9	1.0
	0.8	4.0
plysμA-Δ4	1.9	1.0
	0.3	4.0
plysμA-Δ5	0.7	8.8
plys-SV40	0.02	1.0
	N.T. (< 0.002)	+ (plysμ)
plys-SV40-3'H2A$^+$	N.T. (< 0.002)	1.0
	N.T. (< 0.002)	+ (plysμ)
plys-SV40-3'H2A$^-$	N.T. (< 0.002)	1.0
	N.T. (< 0.002)	+ (plysμ)

N.B. See table legend associated with table 5.1.

frequency, contradicts the result obtained with plysμ2. Hence, this result suggests that transcriptional interference is not the reason why some two-gene plasmids give low, stable *gpt*$^+$-transfection frequencies.

5.5. Summary and discussion.

This chapter describes results (primarily, stable *gpt*$^+$-transfection frequencies) which suggested that *gpt* transcription can be severely restricted, in two-gene expression plasmids containing divergent transcription-units (which, therefore, abut at their 3'-ends). A hypothesis involving transcriptional interference was proposed to explain this phenomenon; it was envisaged that RNA-polymerase, after transcribing the non-selected-gene sequence, continues transcribing into the 3'-end of the *gpt* transcription-unit, thereby restricting the capability of RNA-polymerase on the *gpt* transcription-unit to produce complete *gpt* transcripts. In contrast, there was no reproducible evidence from stable *gpt*$^+$-transfection frequencies (not shown), that the ΔTK promoter of the *gpt* transcription-unit in the two-gene tandem expression-plasmids (see section 4.8. and figures 4.19. and 4.20.), which is downstream of a non-selectable gene transcription-unit, was greatly suppressed by the type of promoter-occlusion phenomenon described in section 5.1.

The motivation for studying the possibility of transcriptional interference was that if sequences which overcome transcriptional interference on *gpt* transcription could be identified, they would most likely be applicable to use in all expression plasmids for J558L. Transcriptional interference on lysozyme transcription-units was not assayed for, but it was proposed that transcriptional interference could be a two-way process. This possibility was suggested by the production of lysozyme transcripts by plys$^-$(d) transfectants, which might be produced by read-through of RNA-polymerase from the *gpt* transcription-unit into the lysozyme coding-sequence. It was reasoned that the inclusion of termination sequences downstream of the selectable-gene transcription unit might overcome any transcriptional interference on the non-selected gene, which might result in high-level expression of the protein of interest on a routine basis.

The finding that plysμA-Δ5 gives efficient expression from the

gpt gene seems to contradict the working hypothesis that *gpt* transcription is only efficient when RNA-polymerase transcribing the non-selected gene is prevented from extending into the 3'-end of the *gpt* transcription-unit; in this plasmid, RNA-polymerase is expected to transcribe into the 3'-end of the *gpt* transcription-unit from the lysozyme coding-sequence. RNA analysis (nuclear run-off) is required to show that plysμA-Δ5 does, indeed, generate a large number of lysozyme transcripts, which would cause the potential for interference on *gpt* transcription. It would be surprising if this is not so, since transcription of the lysozyme coding-sequence is driven by the IgH-gene enhancer and promoter. It is true that all two-gene plasmids with transcription units which abut at their 3'-ends do give slightly lower, stable gpt^+-transfection frequencies than pPROM-Egpt constructs, which may suggest that some transcriptional interference does take place. Alternatively, as discussed in section 5.2.1., this may be because the enhancer is more distant from the *gpt* transcription-unit. In conclusion, transcriptional interference seems to be, at most, of only minor importance in determining the level of expression from the *gpt* gene in two-gene plasmids. Therefore, the question of why some of the two-gene expression plasmids give such poor expression from the *gpt* gene remains unanswered.

There are large differences in the sequences and the relative positioning of the transcription units in some of the plasmids whose stable gpt^+-transfection frequencies were compared in this chapter. Therefore, the expression level from the *gpt* genes may be different due to non-specific influences rather than through specific, controlled mechanisms. For example, the average copy-number of integrated plasmids, in transfectants containing different plasmids, may show variation. A more likely explanation, is that the relative, spatial positioning of the two promoters, combined with effects of other sequences in the plasmids, causes different conformational arrangements of the chromatin within and flanking the transcription units of the integrated plasmids. This might result in different transcription-initiation rates of the (*gpt*) gene in different plasmids. The effect of conformation on expression was discussed in section 5.1. Unfortunately, because the current understanding of the conformational-factors affecting transcription rates is very basic, it is not possible to predict specific mechanisms to explain the different *gpt*-transcription rates on the different plasmids. Nuclear 'run-off'

transcriptional analysis on isolated nuclei from transfectants containing different plasmids is required to compare transcription-initiation rates, to investigate if the different *gpt*$^{+}$-transfection frequencies might, indeed, be due to differences in the rate of transcription initiation.

It might be argued that the results from this chapter cast doubt on the reliability of stable-transfection-frequency assays for comparing the activity of various expression elements. However, the assay seems more reliable when comparing the effect of small, localised changes in expression plasmids, such as in the expression plasmids compared in chapter 3. In that series of experiments, the reproducible trends obtained with different combinations of upstream expression-elements demonstrated the importance of the IgH enhancer; this is a finding which agrees with the findings of numerous other groups using alternative assays. Also, this information proved useful, as it was used in the design of expression plasmids which were then shown to be capable of giving high-level expression. This seems to justify the strategy used. Even so, stable-transfection-frequency assays are rather crude and prone to misleading conclusions, because they are an indirect measurement and rely on many assumptions (see sections 3.1. and 3.2.). However, direct measurements on stable transfectants, such as estimation of the transcription rate by nuclear 'run-off' experiments, or measurement of stable-mRNA levels on Northern blots, in order to estimate the activity of expression elements, are also subject to misleading results due to the non-specific conformational effects. In addition, as described in section 3.1., such analysis necessitates the use of a pre-selected and potentially unrepresentative cell population. This seems to swing the pendulum back in favour of transient assays but, as described in section 3.1., these may also give misleading conclusions.

In conclusion, stable-transfection-frequency assays should, usually, only be used in preliminary screening of large numbers of expression elements, in order to identify potentially-efficient elements, whose efficiencies should then be confirmed by alternative assays.

CHAPTER 6: GENERAL DISCUSSION AND FUTURE PROSPECTS.

This project involved the initial stages of the development of the mouse myeloma cell-line, J558L, as an expression host for proteins encoded by heterologous genes. Expression plasmids were constructed which can give rise to J558L transfectants secreting recombinant lysozyme, at levels typical of the level of secretion of Ig chains in myeloma cells, encoded by endogenous Ig-genes. Expression from a transfected Ig-gene was also obtained at a similar level. In this chapter, I discuss how the expression system might be further developed to increase its efficiency. Also, I wish to emphasise that the use of mammalian expression-systems is not restricted only to the high-level production of proteins with potential therapeutic uses; they may also be used as a research tool, for cloning of genes, and for academic studies to further the understanding of cellular metabolism, signalling and function.

At the beginning of this project, it was expected that both lysozyme and Ig, encoded by transfected genes, could be secreted from J558L at high levels, providing mature mRNAs encoding these proteins could be efficiently produced and then translated. This is because post-translational considerations, such as folding and chemical modifications, which might influence the rate of passage through the secretory pathway (see section 1.9.7) and, hence, the level of protein secretion, are likely to be minimal, compared with the influence on other heterologous proteins; lysozyme is a simple protein, which appears to undergo limited post-translational modifications (see section 4.1.), and is normally secreted; and myeloma cells are specifically geared to expression and secretion of Ig (see section 1.12.). Even so, as alluded to above, the secretion of high levels of these transfected-gene products requires high-level mRNA production. This, in turn, suggests that the transfected genes encoding these proteins are efficiently transcribed in some of the transfectants, such as clone D8 (see section 4.7.6.) and clone D4 (see section 4.3.). Hence, the transfection / expression system used is likely to also be capable of giving rise to a high-level transcription of other transfected coding-sequences and, therefore, at least give the potential for high-level expression of the encoded heterologous proteins. However, the system does

not yet lend itself to the routine isolation of transfectants expressing high levels of the products encoded by transfected genes. There are several ways in which the efficiency of the system could be improved, as outlined below.

One strategy for increasing expression levels is to increase the gene copy-number of transfected sequences. The EBV-based expression system (see section 1.5.1.3.), in which transfected plasmids are maintained episomally, may be worth investigating for this purpose. In this system, the plasmids are reported to be maintained at an average copy number of 10 per cell. This is greater than the copy number sometimes obtained with stably-integrated passively-replicated expression plasmids (e.g. Mulligan and Berg, 1981a); this is the type of expression vector that was used in this study. However, the average copy-number of integrated pova(d) sequences in pova(d) transfectants of J558L, as determined by Southern-blotting analysis (see figure 4.11.), seems to be at least as high as the copy number of the above described EBV-based expression plasmids. If this is the case for all two-gene plasmids, an integrated system may actually be more suited to giving a high, average copy-number. This is especially so if the integrated expression plasmid sequences can be amplified (see section 1.8.). However, attempts by other workers to obtain an amplification system based on the selectable, *E. coli* gene, *gpt*, which was used in this study, have been unsuccessful. As described in section 3.1. and illustrated in figure 3.1., this selection system depends on expression of the enzyme, XGPRT, from the transfected *gpt*-gene. XGPRT catalyses a step in the biosynthesis of guanine nucleotides via a 'salvage' pathway, in medium (XHMPA supplemented) which inhibits endogenous guanine-nucleotide biosynthesis. Ideally, to obtain amplification of the *gpt* gene (and associated flanking sequences), a stoichiometric inhibitor of XGPRT would be used. Unfortunately, no such inhibitor of XGPRT is currently available. Hence, some other strategy is required to obtain amplification of the *gpt* gene. Other work in the laboratory (T.M. Harrison, personal communication) has shown that the fed concentration of xanthine, which is the substrate used (salvaged) by XGPRT in XHMPA selective-medium, can be decreased gradually by two orders of magnitude in selective medium with no detectable effect on growth of stable gpt^+-transfectants. If the xanthine concentration becomes limiting, it might allow for selection of cells that

have undergone amplification of the *gpt* gene, because such cells might direct a greater proportion of the xanthine towards guanine-nucleotide biosynthesis. Of course, for such a system to give satisfactory levels of amplification requires that XGPRT has a poor affinity for xanthine. If the Km of the enzyme is low (i.e. has a high affinity for xanthine), the concentration of xanthine may not be limiting, even at very low concentrations of xanthine. Hence, in this situation there would be no selection for amplification of the *gpt* gene. Also, even if this approach was theoretically feasible, the low concentration of xanthine in the medium might result in it being rapidly exhausted. Hence a perfusion (continuous feed) system would probably be required to replenish the xanthine.

Another selection system commonly used in myelomas is that based on the *E. coli* gene, *neo*, which encodes the enzyme, neomycin phosphotransferase (e.g. Hendricks *et al.*, 1988). This enzyme confers resistance to the neomycin analogue, G418, which is used in the selective medium. Again, there is no stoichiometric inhibitor of this enzyme available to allow selection for gene amplification. The *neo* gene might also be amplifiable through a gradual increase in the concentration of G418 added to the medium. For such a strategy to be successful, it requires that the V_{max} of the enzyme is quite low. Otherwise the turnover rate of the enzyme might continually increase as the concentration of its substrate, G418, is increased, resulting in no selection for gene amplification. In addition, the concentration of G418 used in the selection itself is approximately 600µg/ml. Hence, it may precipitate out of solution at concentrations required to cause large-scale amplification.

As the selection systems presently used for myelomas are not readily amplifiable, it may be necessary to explore the use of alternative selection / amplification systems. As discussed in section 1.7., there is an increasing number of selection systems available. Even the systems which select for expression from a transfected eucaryotic-gene, even though there may already be endogenous low-level expression from the same gene in the host cell, are amplifiable in some cell-types. It will be necessary to empirically screen these various selection systems to determine which one is the most appropriate to J558L. In general, however, amplification is more likely to be efficient on selectable genes which are not expressed endogenously. In the case of eucaryotic selectable-genes, such a selection

system is likely to work on the basis of complementation of an auxotrophic mutation in the host cell. As discussed in sections 1.7. and 1.8., the DHFR selection / amplification system can be highly efficient in DHFR$^-$ CHO-cells. Unfortunately, there is no DHFR$^-$ myeloma-cell-line available at present, and it is probable that the modified versions of the DHFR amplification-system (see section 1.8.), which allow DHFR$^+$ cell-lines to be used, will prove equally unreliable in myeloma cells as they have in other cell-types. Hence, a DHFR$^-$ myeloma-cell-line may be worth screening for. Another potential selection / amplification system for use in myeloma cells is one based on the glutamine-synthetase gene. Glutamine is added to the growth medium of myeloma-cell cultures to allow good growth and, therefore, a selection system for glutamine synthetase is likely to be relatively simple to develop. Again, it remains to be determined if this gene is amplifiable in myeloma cells. Another selection system which might be suitable for use in myeloma cells is that based on the HGPRT (hypoxanthine-guanine phosphoribosyltransferase) gene, which is used in combination with HAT-containing selective medium (see section 1.7.). HGPRT$^-$ myeloma-cell-lines are available which do not readily revert. Such cells are widely used in the generation of hybridomas for monoclonal-antibody production; they are fused with HGPRT$^+$ lymphocytes, and immortal HGPRT$^+$-hybridomas are selected for in HAT-supplemented medium. There are also several inhibitors of HGPRT available, which may allow for amplification of the HGPRT gene. It is possible that a HGPRT$^-$ derivative of J558L could easily be produced, using the same methods as used in the past for isolation of other HGPRT$^-$ myelomas.

This study, as have many previous studies, showed that the expression level of a heterologous protein in transfectants containing the same expression plasmid is highly variable (e.g. see figures 4.5. and 4.25.). Although this may be partially explained by variation in the copy number of integrated sequences (see section 1.8.), another major reason is likely to be the site of chromosomal integration (see section 1.9.1.3.). The use of an expression vector which is episomally maintained (see section 1.5.) is likely to result in less variation in expression levels between transfectants. This is because all the transfected vectors are likely to be in similar environments and, therefore, comparative expression-levels from different transfectants should depend primarily on copy number.

As discussed in section 1.9.1.3., Grosveld *et al.* (1987) showed that the inclusion, in expression vectors, of DNaseI-superhypersensitive sequences normally associated with the mouse globin-gene locus, allows for position-independent expression from integrated transfected sequences in transgenic mice. Hence, in the presence of such sequences, the level of transcription from one β-globin - mini-gene is similar to the normal levels of native β-globin - gene transcription, and for higher copy-numbers of the gene, the level of transcription increases directly in proportion. The effect of these globin-gene-locus sequences seems to be cell specific and has only been identified in erythroid cells. Consequently, they are likely to be of no use in myeloma cells. There is speculation, however, that there are similar sequences associated with many gene loci. It is believed that they are more likely to be associated with genes whose expression is tightly controlled in a spatial and temporal manner; when such genes are expressed, they are generally expressed at particularly high-levels, and so it logical to speculate that their expression is caused by elements in addition to an enhancers and promoters. It is thought that DNaseI-superhypersensitive sites are less likely to be associated with housekeeping genes, which are expressed in all or most cell-types at similar levels, and encode the proteins performing functions common to most cells.

In myeloma cells, the gene loci associated with Ig expression are the most likely to contain active sequences with the same properties as those described by Grosveld and co-workers. The identification of such a sequence would surely make myeloma cells one of the most favoured mammalian expression-hosts. The fact that large parts of gene loci associated with Ig expression have been cloned may make this possible, before similar sequences with specificity towards alternative expression-hosts are identified. The IgH-gene locus is, however, large and complex, and is not well defined in the 5'(V_H)-regions. It may make more sense to test DNase-hypersensitive sites found in smaller genes for their ability to confer position-independent expression. For example, the J-chain gene encodes a protein which is necessary for pentamer IgM assembly. The level of J-chain expression increases during B-lymphocyte development in tandem with the increased expression of IgH(μ) and IgL chains (Roth *et al.*, 1979), and this is concurrent with the appearance of DNase-hypersensitive sites in the J-

chain gene (Minie and Koshland, 1986). In addition, a sequence element within this DNase-hypersensitive region is complementary to elements known to be involved in control of expression from the IgL(κ) gene (Falkner and Zachau, 1984). Perhaps this sequence, possibly combined with other similar sequences, could enable position-independent expression of transfected genes in myeloma cells. Another gene which is highly expressed in myelomas, and in fact in most cells, is that encoding β-actin. Although it might be considered a housekeeping gene because of its widespread expression, it is possible that because of the high level of expression from this gene, there may be transcription activating elements additional to the promoter. In fact, expression vectors have been constructed containing the β-actin - gene promoter and 5'-flanking region, and these drive high-level heterologous-gene expression in a large number of cell-lines (Gunning *et al.*, 1987). If a sequence from the β-actin gene was found to confer position-independent expression, it might be applicable to a wide range of mammalian expression-hosts.

It would be interesting and informative to gain a further understanding of why the plys^{+}(d) and plys-SV40 transfectants of J558L, clones D8 and C6 respectively (see section 4.7.6.), express such high levels of lysozyme. As discussed above, the variation in expression from a transfected gene in transfectants containing the same expression vector is due to a combination of gene copy-number and chromosomal location. Southern-blotting analysis needs to be performed on these clones to investigate if they have a particularly high copy-number of plasmid sequence. More extensive Southern-blotting analysis might also give some indication of the arrangement of the plasmid sequence(s) in the genome. If such analysis did not reveal anything distinctive about these clones, it might suggest the plasmids are integrated into a chromosomal locus favourable to high-level transcription. The transfected plasmid sequences could be recovered along with some flanking genomic sequences, by genomic cloning. Then further genomic flanking sequences could be cloned using the strategy of chromosome walking. Such genomic clones might lead to identification of an important transcription-control sequence, perhaps similar to the 'Grosveld' sequences, with activity in myeloma cells. Inclusion of such a sequence in expression plasmids might allow routine high-level transcription from transfected genes in myeloma cells. It might

be the case that genomic cloning only results in partial recovery of sequences required to allow position-independent transcription. However, it is likely that selection systems for targetted homologous-recombination (see section 1.6.) will improve in the near future. Therefore, inclusion of some of the cloned genomic-sequences in a plasmid might allow its targetting by homologous recombination into the locus containing the remaining sequences which confer position-independent transcription.

It has to be emphasised that 'Grosveld-like' sequences might not necessarily increase the level of protein expression obtainable above, for example, the level of lysozyme expression from the plys^{+}(d) transfectant, clone D8. This is because of limitations on the level of expression of any protein, which are discussed in detail later in this chapter. For mammalian-cell expression systems in culture, the advantage of vectors containing such sequences is that a greater percentage of transfectants would generate high levels of transcripts, and so yield the maximum level of protein expression. This would reduce the amount of screening needed to isolate high-level expressing clones. Of course, these sequences are of even greater potential value for the purpose of gene therapy, where tightly controlled expression, and usually high-level expression, from transfected genes will be essential.

Attempts to explain the apparent interference between transcription units in some two-gene plasmids were unsuccessful (see chapter 5). Therefore, this phenomenon remains a potential problem which may limit expression from both genes in an unpredictable manner. As discussed in section 3.1., if transfection procedures continue to improve, it may become possible to transfect any gene and obtain stable expression at a sufficiently high frequency, such that a selectable-marker gene for stable transfectants is no longer required. This would allow expression plasmids containing only a single gene, encoding the protein of interest, to be used. It might be argued that the absence of the selectable-marker gene is undesirable, because the retention of transfected sequences can no longer be selected for. However, preliminary studies (see section 4.4.) suggest that transfected sequences in J558L are reasonably stable in the absence of selection. Even so, there may be other advantages in using a selectable-marker gene. For example, amplification systems are normally based on the selectable gene and, therefore, amplification of the

transfected sequences would not be possible using the above approach.

If the observation of section 4.7.6., that some slow-growing colonies (during selection in XHMPA selective-medium) give the highest expression from the non-selected gene, is reproducible, it may be worthwhile to rescue slow-growing colonies 2-2½ weeks post-transfection. This might allow cells which express XGPRT at low levels (just enough to survive selection) to be rescued, which might otherwise perish by the time colonies are routinely picked (21 days post-transfection). It is most likely that some of these clones express low levels of XGPRT because the site of integration of the expression plasmid into the chromosome is not the most favourable for transcription and/or the copy number of the plasmid sequences is low. However, as speculated in section 4.7.6., the low level of XGPRT expression in some of these colonies might also be due to high-level transcription from the lysozyme transcription unit. For example, as discussed in chapter 5, the chromatin-structure associated with the active lysozyme-transcription-unit(s) might interfere with the *gpt* transcription unit(s). Alternatively, as discussed later in this chapter, the high level of lysozyme expression might be a drain on the cell's resources, making it less likely to survive the selection process.

A problem with the above strategy is that the XHMPA selective-medium kills non-transfected cells at a rather slow rate; it takes over a week before the non-transfected cells stop growing completely. There are two possible reasons for this. The first reason, and one which applies to all selection systems, is that during initial stages of selection many cells contain transiently-existent extrachromosomal plasmids (see section 1.6.), which can encode resistance against the selective agent(s). But the fact that the *gpt* selection-system takes longer to take effect than most other systems, suggests that even as the majority of cells lose the transfected plasmids, guanine nucleotides are still produced. This might be because mycophenolic acid only partially blocks the endogenous pathway leading to guanine-nucleotide biosynthesis (see figure 3.1.) or, alternatively, there is some other pathway leading to guanine nucleotide biosynthesis. As a consequence of the selection being slow to take effect, if colonies are picked soon after the transfection experiment, one runs the risk of picking non-transfected cells. Even so, the risk might be worthwhile if it allowed isolation of particularly high-level expressing

transfectants. Alternatively, this risk might be overcome by adding HAT (hypoxanthine, aminopterin and thymidine) to the XHMPA selective-medium (Mulligan and Berg, 1981). The aminopterin inhibits several steps in the *de novo* pathways of nucleotide biosynthesis, and this can speed up the selection process. This might result in death of all non-transfected cells before cells expressing low levels of XGPRT begin to die.

All the above modifications that might be made to the expression system are likely to increase the level of transcription. However, as previously discussed, the isolation of transfectants expressing high levels of heterologous proteins (e.g. clone D8), strongly suggests that high-level transcription is already obtainable with the vector/transfection/selection system developed. The capability of a cell to produce and secrete large amounts of functional heterologous-protein following high-level transcription of the encoding gene, is likely to be an intrinsic property of the cell which is independent of the type of expression system used. As discussed in section 1.3., this is why it is desirable to have an expression system (or systems) available for several different host-cell types, so that their suitability to produce, and preferably secrete, a particular protein may be compared.

Other members of the laboratory (M. Chidgey, personal communication) have utilised modified versions of the divergent two-gene expression plasmids (see section 4.7.) to express heterologous proteins, larger and more complex than lysozyme. Tissue-plasminogen activator has been secreted at levels up to approximately 0.4μg/10^6 cells/24 hours, which is comparable with the optimum expression-level of tissue-plasminogen activator obtained from J558L by Weidle and Buckel (1987). Human pro-renin can also be expressed and secreted, although, as yet, only low-level expression (an optimum level of ≈ 0.1μg/10^6 cells/24 hours) has been obtained. It seems that J558L is unable to proteolytically cleave the 'pro' sequence from pro-renin, prior to secretion (interestingly, renin can be secreted in the absence of a 'pro' sequence; paper in preparation). The secretion of pro-renin shows that J558L is able to secrete (an alternative form of) a protein which is normally secreted by a regulated rather than a constitutive secretory-pathway (see section 1.11.). This suggests that a protein can be packaged into and secreted from a non-clathrin-coated secretory vesicle, when there is not the alternative of a nascent storage-

vesicle to be packaged into. Hence, J558L may be a suitable secretory-host, providing the levels of secretion are satisfactory, for routine secretion of all classes of secretory protein. However, if a proteolytic cleavage is normally necessary within storage vesicles prior to secretion in order to generate the active protein, it seems likely that some means of performing the cleavage, following expression and secretion from J558L, will be essential. For, pro-renin, this is easily performed by addition of trypsin.

Both the tissue-plasminogen activator and renin (generated by trypsin treatment of pro-renin), produced in J558L, possess biological activity. This, along with the demonstration that lysozyme produced in J558L is also active (see section 4.7.6.), suggests that myeloma cells are capable of performing whatever (if any) post-translational modifications (see section 1.10.) are necessary for functional activity. Yet more proteins will have to be expressed to assess the true versatility of J558L as an expression host.

The production of a heterologous protein at high levels is likely to impose a considerable burden on the resources of the cell. Such a protein usually plays no role in the growth or well-being of the cell, yet its production consumes a large proportion of the amino-acid pool and energy store. Therefore, if the expression vector is particularly efficient, a situation may occur in which the potential for expression of the heterologous protein is so high, that it is detrimental to the growth of the cell. Hence, there is a limit on the level of expression of a heterologous (or homologous) protein that can be obtained. It is difficult to predict what this limit will be, and it will have to be empirically determined for each protein and cell-type. It is possible that the reason why clones C6 and D8 grew particularly slowly in XHMPA-supplemented medium during selection (see section 4.7.6.), is that their high level of lysozyme expression is detrimental to cell growth. This is likely to be particularly apparent during the selection process, when the vitality of the cells is not promoted to any great extent by cross-feeding from neighbouring cells, in contrast to normal cell-culture in which the cells are present at a much higher density.

A cell may require particularly nutritious medium to allow optimum expression of a heterologous protein, but because of the high cost of cell-culture growth medium, a balance has to be struck between product

yield and the cost of producing it. Clearly, the further development of serum-free medium and its optimisation for each cell-type, will be as important as optimisation of expression vectors, transfection systems and selection/amplification systems, in obtaining cost-effective high-level expression.

It is often the case in cell-culture systems, that production of a desired product is incompatible with cell-growth. Therefore, a medium could initially be used which allows accumulation of cells to high density. Following establishment of the culture at high cell-density, the medium could be changed to one which is not so supportive of cell division, yet allows the desired protein to be produced. The use of inducible expression-elements may also be of use for overcoming the incompatibility of cell growth and heterologous-protein production. These strategies are likely to be particularly important when producing proteins which are cytotoxic. If such proteins could be produced constitutively, one might never obtain transfectants because of the toxicity. In contrast, with a tightly controlled-expression system, cells can be cultured to high density and good condition, before switching on expression. The expression might be switched on long enough to allow the expression product to be harvested, yet short enough so that all cells are not killed. As described in section 3.7., induction of the mouse metallothionein-I promoter or the MMTV-LTR could not be demonstrated in J558L transfectants. An alternative, inducible promoter which is active in mammalian cells (CHO cells) is the *Drosophila hsp*70 promoter (Wurm *et al.*, 1986). This promoter is induced at 43°C, but is inactive at 37°C. Hence, transcription can be induced by incubating an established culture at 43°C for a short period, and then the culture can be incubated at 37°C, to allow recovery, and translation of the mRNA generated from the induced transcription unit. The activity of this promoter in J558L needs to be established; for tight control of expression the promoter would probably have to be used in the absence of the IgH enhancer, but as seen in section 3.4., some promoters appear to have poor activity in J558L in the absence of the IgH enhancer.

At the outset of this study, using spheroplast fusion, a J558L cell was the most-easily-transfected myeloma cell. This was the major reason for the use of the J558L cell-line. However, it may not be the perfect myeloma expression-host. Although it does not produce an IgH chain,

it still produces and secretes a λ light-chain (see figure 4.16.). The production of the λ chain accounts for a significant proportion of the cell's resources. Ideally, a non-Ig-producer host cell-line would be used, which might result in some of the resources 'wasted' in J558L, on production of λ chain, being redirected towards increased production of a heterologous protein. In addition, although myeloma cells have an extensive secretory-pathway-network (see section 1.11.), it may only be large enough to serve the normal functions of Ig secretion and targetting of intracellular traffic. Hence, if a heterologous protein is expressed in an Ig-producing host cell, the level of heterologous-protein secretion may be restricted due to competition for limited 'space' in the secretory pathway. One way of overcoming these potential problems, is to screen for a derivative cell-line of J558L which has lost the ability, preferably at the level of the gene (e.g. a deletion of transcription-activating elements), to express λ chain. This could be achieved by large-scale subcloning and then screening for loss of secretion using some simple assay, such as an ELISA, followed by an investigation at the RNA level to identify cells not generating transcripts from the λ gene. Alternatively, a non-Ig-producer myeloma cell-line which is presently available could be used; it is highly likely that the expression-plasmids developed for J558L are also suitable for use in these cell-lines. One potential candidate is the non-Ig-producing mouse-myeloma cell line, NSO (Galfrè and Milstein, 1981). Attempts to obtain stable *gpt*$^+$-transfectants by transfection of NSO with pMMTV-Egpt (see section 3.3.1.), using spheroplast fusion, resulted in the generation of only one transfectant from two transfection-experiments. Hence, this cell line appears to be resistant to transfection by spheroplast fusion, especially when it is considered that pMMTV-Egpt was seen to be the most efficient of the *gpt* expression-plasmids analysed in this study (see section 3.3.2.). However, other work in the laboratory (S. Munson, personal communication) has demonstrated that NSO is much more amenable to transfection by electroporation (see section 1.4.). Hence, the level of expression of heterologous proteins from this cell line can now be compared with that from J558L.

The biochemical properties of the J558L cell-line could also be manipulated to make it a more cost-effective expression host. It has already been alluded to that glutamine has to be included in the growth

medium of a J558L culture. This need could easily be alleviated by transfecting J558L with an expression plasmid containing the glutamine-synthetase gene. It is likely that a high proportion of the transfectants would have the desired property of glutamine-independent growth because optimum expression is unlikely to be required to confer this property. Another exciting prospect is the potential use of anti-sense RNA to redirect product flow and stop processes detrimental to product generation. Anti-sense RNA is RNA which is complementary to a given mRNA species, with which it pairs, forming double-stranded RNA. In one way or another (reviewed by Green *et al.*, 1986; Lichtenstein, 1988), duplex-RNA formation blocks expression of the encoded protein. A sequence, encoding anti-sense RNA, could be transfected into J558L on one of the expression plasmids described in this study, so as to generate anti-sense transcripts. It must be remembered, however, that every time a different gene is transfected into J558L, a plasmid carrying a different selectable marker is required. Therefore, the extent to which this approach of altering the host-cell's metabolism can be used, is limited by the number of selectable-marker systems that are currently in use with myelomas. This again illustrates the need to apply and optimise some of the selection systems described in section 1.7. to their use in myeloma cells. Another restriction on the manipulation of the biochemical properties of an expression host is the lack of basic knowledge on the importance of some biochemical pathways. This may make it difficult to rationally identify characteristics of the cell, which when changed, would improve the efficiency of the expression host.

The preceding discussion has emphasised that a considerable improvement of the expression system developed in this study is required, before J558L could be commercially exploited as an expression host for production of therapeutic proteins. However, the system might be applied to basic-research purposes as discussed below. These studies are unlikely to require optimum levels of heterologous-protein expression for them to be of use.

One potential application of a mammalian expression-system is for the cloning of eucaryotic genes as cDNAs. There are, in fact, several strategies for cloning a cDNA which do not involve the use of a mammalian expression-system. For those genes that give rise to mRNA of low abundance

and which, therefore, cannot be cloned directly, one of the following two approaches, not involving a mammalian expression-system, is most commonly used. The first approach can be used when the protein encoded by the gene has been purified. This may allow partial amino-acid - sequence data to be determined. Oligonucleotide probe(s) based upon a peptide amino-acid - sequence can then be synthesised and used to screen an appropriate gene-library for hybridisation to the complementary sequence (Montgomery *et al.*, 1978). The second approach is to express part of the protein of interest, encoded by a partial-cDNA clone, in the form of a β-galactosidase-fusion protein in *E. coli*. This can then be screened for amongst recombinant clones with an antibody, specific for the protein of interest (Young and Davis, 1983). Both these approaches require prior information and/or materials to allow isolation of the cDNA; in the former approach, the protein sequence may be difficult to obtain because, for example, of difficulty in purifying the protein, or because the sequence is 'blocked' at its N-terminus; the latter approach requires a pre-existing monospecific antibody, and that the antigenic determinants are recognised in the context of the fusion-protein. Therefore, neither strategy is applicable to the cloning of all genes.

In contrast to the above strategies, the approach of 'functional mammalian-expression cloning' relies solely upon expression from a cDNA copy of the gene of interest in mammalian cells (Okayama and Berg, 1983). Therefore, unlike the above described methods, there is no requirement for purified protein or antibody; in fact, no details of the protein are required at all. A library of cDNAs could be cloned into one of the expression plasmids constructed in this study, and then transfected into a population of J558L cells. During the development of functional mammalian-expression as a technique for cloning cDNAs, transfected cells carrying a cDNA of interest were usually identified by complementation of a mutant defect in the expression host. If using J558L as the expression host for expression cloning, however, transfectants carrying the cDNA are more likely to be identified by an alteration in the cell's morphology or growth phenotype, or through the cell's expression against a negligible background of a binding activity or some other readily assayable function.

It must be emphasised that the functional - mammalian-expression approach to cDNA cloning is not always straightforward. Firstly, it

requires the technical ability to synthesise cDNAs of full length, so as to encode a functional protein. This, of course, is independent of the expression system. Also, in order to isolate cDNAs corresponding to low-abundance mRNAs, a high transfection-frequency is required so as to obtain sufficient transfectants to provide a full representation of the cDNA library. Transfection frequencies routinely obtained through spheroplast-fusion transfection of J558L would not satisfy this requirement, particularly with the divergent two-gene plasmids (see section 4.7. and table 5.1.). It is usually the case that a selection for stable transfectants is made before screening for the clone carrying the cDNA of interest. This necessitates the use of a two-gene expression vector. Perhaps the use of tandem two-gene expression plasmids (see section 4.8.), which gave higher transfection frequencies than the divergent plasmids (usually 10^{-4} - 10^{-5} transfectants per cell), combined with the optimisation of J558L-transfection using electroporation, would allow sufficient stable transfectants to be generated. Another potential problem is that the expression plasmids used in this study depend on integration into the host genome for their propagation (see section 1.5.1.1.). This presents the problem of recovering the cDNA once a transfectant, expressing the protein encoded by the cDNA of interest, has been identified. One method that can be used is to include a SV40 replicon in the transfected plasmid. Then the transfectant carrying the required cDNA is fused with a COS cell. As described in section 1.5.1.1., COS cells express the large T-antigen of SV40 constitutively, and this promotes replication of sequences containing SV40 replicons. Hence, in the hybrid cell, it promotes replication, and consequently excision of the integrated plasmid-sequences as extrachromosomal, circular DNA. These can then be isolated and cloned, following transformation into bacteria. In contrast, the use of episomally-maintained expression vectors obviates the need to excise transfected sequences from the chromosome. Therefore, they are much easier to transfer from a mammalian cell to a bacterial cell. For this reason, cloning expression-plasmids based on EBV, described recently by Margolskee *et al.* (1988) (see section 1.5.1.2.), may be preferable to the plasmids used in this study, for cloning by functional expression.

One of the most difficult sets of genes to clone at present are those encoding the *trans*-acting transcription-regulating proteins. As

discussed in sections 1.9.1.1. and 1.9.1.2., these interact with *cis*-acting sequences normally in the 5'-flanking region of a coding sequence, so as to determine the transcription rate. It is difficult to purify the factors because they are present at low levels and this has restricted, particularly for mammalian-cell factors, the cloning of their genes. However, the availability of efficient cloning-systems based on functional mammalian-expression should allow more rapid cloning of the genes. For example, the introduction of a cDNA encoding a *trans*-acting factor into a cell might result in activation of a gene, which is normally inactive because the *trans*-acting factor is not normally expressed in the cell.

Traditionally, basic research on cellular processes has been performed using biopsy or autopsy animal-tissue. This approach is becoming increasingly unacceptable because of pressure from animal rights organisations. In addition, there are not always ideal or readily-available animal models for the investigation of human diseases. Consequently, cell-culture systems are gaining wider acceptance for use in academic studies on cellular processes and in the preliminary stages of screening for various therapeutic and pharmacological agents. Combined with the potential to express exogenous genes in such cells, they may become more widely used. Below, are discussed two potential uses of mammalian expression-systems for the study of fundamental cell processes.

As discussed in section 1.11., myeloma cells have a highly-developed secretory-pathway network to allow high-level Ig secretion. Hence, they might be used as a model system for studying the secretory pathway of mammalian cells. The components of the secretory pathway in mammalian cells are well defined (see section 1.9.7.), but the factors and mechanisms controlling the passage of a protein through the secretory pathway are barely understood. The key to understanding the flow of traffic is to elucidate the working parts of the system which incorporate specific proteins into specific transport-vesicles. There are two basic approaches being used (reviewed by Pfeffer and Rothmanm, 1987). One involves biochemical-reconstitution studies, and the second involves genetic studies. So far genetic studies of eucaryotic secretion have primarily been restricted to yeast cells. The secretory pathway in yeast has been defined genetically by a series of temperature-sensitive mutant strains. At least 23 complementation-groups, and thus gene products, are required for

transport of secretory proteins from the site of synthesis to the cell surface (Novick *et al.*, 1980). These genes are now being cloned and the products they encode are being characterised. Although mammalian cells are not as simple to manipulate as yeasts, the availability of the J558L expression-system means that, in principle, parallel experiments can be performed to those described above for yeasts. For example, a secretory-pathway mutation could be complemented by expression-cloning, and the cDNA isolated could be the first step in the characterisation of an essential component of the secretory pathway. In addition to studying the components of the secretory pathway, it may also be important to identify signals in the proteins themselves. Such sequences are proposed to allow retention in the secretory pathway, or targetting to other cellular locations such as the lysosomes (Wieland *et al.*, 1987). One strategy would be to mutagenise a gene prior to its transfection into the expression host. If the deletion of a sequence resulted in a change in the targetting of a protein, it might allow identification of a sequence necessary for targetting. However, previous attempts using this type of strategy have not led to the isolation of targetting signals. As explained in section 1.9.7., this may be because the signals are composed of various regions of the protein, that are only brought together during protein folding and are, therefore, conformation dependent. Moreover, an alteration in the targetting of a protein in such a study might be due to non-specific changes in the conformation of the protein leading to the formation of new, cryptic targetting-signals

Another potential use of mammalian expression-systems is the expression of cell-surface receptors for their study in a 'normal' membranous environment. Advantages of studying receptors following production by recombinant means are likely to be numerous. The protein could be expressed at higher levels and, therefore, its effect on the cell following binding of ligand might be greater than in the native cell, which might make its effect easier to study. Secondly, in a similar way to that described above for studying protein-targetting, the expression of mutated receptor-genes may allow identification of the role the different domains or subunits of a receptor play in receptor functions. These include not only the binding function, but also the relay of the message which results in initiation of a cell-response. Thirdly, if the receptor is highly expressed in a host cell, this may allow the isolation of sufficient

material for protein studies which have not been possible or are difficult by conventional means. This would be applicable, for example, to the insulin and type-I - insulin-like - growth-factor receptors which seems to be present at particularly low levels (see Ullrich and Ramachandran, 1987). The increased information and understanding of receptor action obtained through such studies is likely to allow the design of pharmacological agents on a more rational basis than is normal at present. Possible drawbacks in expressing a receptor in a heterologous host are, firstly, that it might have slightly altered characteristics due to the cell-specificity of post-translational modifications and, secondly, that the host cell may not respond to the same extent or in the same manner as the native cell, following binding of ligand to the receptor. Hence, if an expression system is available for the cell-type in which the receptor is natively produced, the use of this cell-type might be preferable to the use of a heterologous cell-type.

Clearly, the potential for mammalian expression-systems is enormous, and more applications as a research tool will become available as techniques, both at the recombinant-DNA level and at the transfection level become more routine.

CHAPTER 7: REFERENCES

Abrahamsén, L., Moks, T., Nilsson, B. and Uhlen, M. (1986) Secretion of heterologous gene products into the culture medium of *E. coli*. Nucl. Acids Res., 14, p7487-7500.

Adetugbo, K. (1978) Evolution of an immunoglobulin subclass - primary structure of a murine γ_1 chain. J. Biol. Chem., 253, p60-75.

Adhya, S. and Gotterman, M. (1982) Promoter occlusion: transcription through a promoter may inhibit its activity. Cell, 29, p939-944.

Akrigg, A., Wilkinson, G.W.G. and Orman, J.D. (1985) The structure of the major immediate early gene of human cytomegalovirus strain AD169. Virus Res., 2, p107-121.

Alderton, G., Ward, W.H. and Fevold, H.L. (1945) Isolation of lysozyme from egg white. J. Biol. Chem., 157, p43-58.

Aloni, Y. and Hay, N. (1985) Attenuation may regulate gene expression in animal viruses and cells. CRC Crit. Rev. Biochem., 18, p327-383.

Alt, F.W., Bothwell, A.L.M., Knapp, M., Sniden, E., Mather, E. *et al.* (1980) Synthesis of secreted and membrane bound Ig mu heavy chain is directed by mRNA that differ in their 3' ends. Cell, 20, p293-301.

Anderson, J.E. (1983) The effect of steroid hormones on gene transcription. In 'Biological Regulation and Development' (eds. Goldberger, R.F. and Yamamoto, K.R.), 3B, Plenum press, N.Y., USA.

Anderson, S., Gait, M.J., Mayol, L. and Young, I.G. (1980) A short primer for sequencing DNA cloned in the single-stranded phage vector M13mp2. Nucl. Acids Res. 8, p1731-1743.

Ashman, C.R. and Davidson, R.L. (1985) High spontaneous mutation frequency of BPV shuttle vector. Somatic Cell and Mol. Genet., 11, p499-504.

Atchison, M.L. and Perry, R.P. (1986) Tandem kappa Ig promoters are equally active in the presence of the kappa enhancer: implications for models of enhancer function. Cell, 46, p253-262.

Aviv, H. and Leder, P. (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA, 69, p1408-1412.

Bachmann, M., Mayet, W.J., Schröder, H.C., Pfeifer, K., Meyer zum Büschenfelde, K.-H. and Müller, W.E.G. (1986) Association of LA and RO antigens with intracellular structures in HEP-2 carcinoma cells. Proc. Natl. Acad. Sci. USA, 83, p7770-7774.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Expression of a globin gene is enhanced by remote SV40 DNA sequences. Cell, 27, p299-308.

Banerji, J., Olson, L. and Schaffner, W. (1983) A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell, 33, p729-740.

Bebbington, C.R. and Hentschel, C.G. (1987) The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells. In 'DNA Cloning' (ed. Glover, D.), 3, p163-188, IRL Press; Oxford, England and Washington D.C., USA.

Benoist, C. and Chambon, P. (1981) *In vivo* sequence requirements of the SV40 early promoter region. Nature, 290, p304-308.

Bently, D.L. and Groudine, M. (1986) A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. Nature, 321, p702-706.

Bergman, L.W. and Kuehl, W.M. (1979) Formation of intrachain disulphide bond in nascent chain immunoglobulin light chains. J. Biol. Chem., 254, p8868-8876.

Bernstein, A., Berger, S., Huszar, D. and Dick, J. (1985) In 'Genetic Engineering: Principles and Methods' (eds. Setlow, J.K. and Hollaender, A.), 7, p235-261, Plenum Press, N.Y., USA.

Bialy, H. (1987) Recombinant proteins; virtual authenticity. Biotechnology, 5, p883-890.

Biggin, M.D., Gibson, T.J. and Hong, E.F. (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA, 80, p3963-3965.

Bingham, P.M., Chou, T.-M., Mims, I. and Zachar, Z. (1988) On/off regulation of gene expression at the level of splicing. Trends in Genet., 4, p134-138.

Birshtein, B., Preud'Homme, J.-L. and Scharff, D. (1974) Variants of mouse myeloma cells that produce short immunoglobulin chains. Proc. Natl. Acad. Sci. USA, 71, p3478-3482.

Bird, A.P. (1986) CpG-rich islands and the function of DNA methylation. Nature, 321, p209-213.

Blanchard, A. (1987) A study of chromatin heterogeneity using immunological methods. Ph.D. Thesis, University of Nottingham, England.

Blanchin-Roland, S. and Masson, J.-M. (1988) Protein secretion by *E. coli* under the control of a synthetic gene. Presented at Biochem. Soc. meeting, Nottingham, England.

Bole, D.G., Hendershot, L.M. and Kearney, J.F. (1986) Post-translational association of IgH chain with nascent heavy chains in non-secreting and secreting hybridomas. J. Cell Biol., 102, p1558-1566.

Bossy, B., Hall, L.M.C. and Spierer, P. (1984) Genetic activity along 315Kb of the *Drosophila* chromosome. EMBO J., 3, p2537-2541.

Bostock, C.J. and Allshire, R.C. (1986) Comparison of methods for introducing vectors based on bovine papillomavirus-1 into mammalian cells. Somatic Cell and Mol. Genet., 12, p357-366.

Brawerman, G. (1981) The role of the poly(A) sequence in mammalian mRNA. CRC Crit. Rev. Biochem., 10, p1-38.

Brawerman, G. (1987) Determinants of messenger RNA stability. Cell, 48, p5-6.

Breathnach, R. and Chambon, P. (1981) Organisation and expression of eucaryotic split genes coding for proteins. Ann. Rev. Biochem., 50, p349-383.

Breindl, M., Harbers, K. and Jaenisch, R. (1984) Retrovirus-induced lethal mutation in collagen I gene of mice is associated with an altered chromatin structure. Cell, 38, p9-16.

Brenner, D.A., Kato, S., Anderson, R.A., Smigocki, A.C. and Camerini-Otero, R.A. (1984) The recombination and integration of DNAs introduced into mouse L-cells. Cold Spring Harbor Symp. Quant. Biol., 49, p151-160.

Brent, R. and Ptashne, M. (1985) A eucaryotic transcriptional activator bearing the DNA specificity of a procaryotic repressor. Cell, 43, p729-736.

Brinster, R.L., Chen, H.Y., Trumbauer, M., Senear, A.W., Warren, R. and Palmiter, R.D. (1981) Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell, 27, p223-231.

Broach, J.R. (1982) The yeast plasmid 2μ circle. Cell, 28, p203-204.

Brown, A.M.C. and Scott, M.R.D. (1987) Retroviral vectors; in 'DNA Cloning' (Ed. Glover, D.), III, p189-212, IRL Press; Oxford, England and Washington D.C., USA.

Butcher, P. and Trifonov, E.N. (1986) Compilation and analysis of eucaryotic POL II promoter sequences. Nucl. Acids Res., 14, p10009-10026.

Butters, T.D. and Hughes, R.C. (1981) Isolation and characterisation of mosquito cell membrane glycoproteins. Biochem. Biophys. Acta, 640, p655-671.

Campo, M.S., Spandidos D.A., Lang, J. and Wilkie N.M. (1983) Transcriptional control signals in the genome of bovine papillomavirus type I. Nature, 303, p77-80.

Campo, M.S. (1986) Bovine papillomavirus DNA: a eucaryotic cloning vector. In 'DNA Cloning' (Ed. Glover, D.), II, p213-238, IRL Press; Oxford, England and Washington D.C., USA.

Carter, A.D, Felber, B.K., Walling, M.J., Jubier, M.F., Schmidt, C.J. and Hamer, D.H. (1984) Duplicated heavy metal control sequences of the mouse metallothionein-I gene. Proc. Natl. Acad. Sci. USA, 81, p7392-7396.

Cartier, M., Chang, M.-W.M. and Stanners, C.P. (1987) Use of the *E. coli* gene for asparagine synthetase as a selective marker in a shuttle vector capable of dominant transfection and amplification in animal cells. Mol. Cell Biol., 7, p1623-1628.

Catanzaro, D.F., Mullins, J.J. and Morris, B.J. (1983) The biosynthetic pathway of renin in mouse submandibular glands. J. Biol. Chem., 258, p7364-7368.

Caulcott, C.A. and Rhodes, P.M. (1986) Temperature induced synthesis of recombinant proteins. Trend in Biotech., 4, p142-146.

Chambon, P. *et al.* (1984) Promoter elements of genes coding for proteins and modulation of transcription by estrogens and progesterone; Recent Progress in Hormone Research, 40, p1-39.

Chan, S.J., Patzelt, C., Duguid, J.R. *et al.* (1979) Information transfer in normal and abnormal cells. In 'From Gene to Protein' (Eds. Russel, T.R. *et al.*), p361-377. Academic Press, N.Y., USA.

Chandler, V.L., Maler, B.A. and Yamamoto, K.R. (1983) DNA sequences bound specifically by glucocorticoid receptor *in vitro* render a heterologous promoter hormone responsive *in vivo*. Cell, 33, p489-499.

Chang, J.Y.-H.,Pai, R.C., Bennett, W.F. and Bochner, B.R. (1988) Periplasmic secretion of human growth hormone by *E. coli*. Presented at Biochem Soc. meeting, Nottingham, England.

Chia, W. and Rigby, W.J. (1981) Fate of viral DNA in non-permissive cells infected with simian virus 40. Proc. Natl. Acad. Sci. USA, 78, p6638-6642.

Chiang, T.-R. and McConlogue, L. (1988) Amplification and expression of heterologous ornithine decarboxylase in Chinese hamster cells. Mol Cell Biol., 8, p764-769.

Christman J.K., Gerber, M., Price, P.M., Flordellis, C., Edelman, J. and Acs, G. (1982) Amplification of expression of hepatitis B surface antigen in 3T3 cells cotransfected with dominant-acting gene and cloned viral DNA. Proc. Natl. Acad. Sci. USA, 79, p1815-1819.

Chu, G., Hayakawa, H. and Berg, P. (1987) Electroporation for the efficient transfection of mammalian cells with DNA. Nucl. Acids Res., 15, p1311-1326.

Church, G.M., Ephrussi, A., Gilbert, W. and Tonegawa, S. (1985) Cell-type-specificity contacts to Ig enhancers in nuclei. Nature, 313, p798-803.

Citron, B., Falck-Pedersen, E., Salditt-Georgiff, M. and Darnell Jr., J.E. (1984) Transcription termination occurs within a 1000 bp region downstream

of the poly(A) site of the mouse B-globin (major) gene. Nucl. Acids Res., 12, p8723-8731.

Clarke, C., Berenson, J., Goverman, J., Crews, S., Sill, G., Boyer, P. and Calame, J. (1982) An immunoglobulin promoter region is unaltered by DNA rearrangements and somatic mutation during B cell development. Nucl. Acids Res., 10, p7731-7749.

Colbere-Garapin, F., Chousterman, S., Horodniceanu, F., Kourilsky, P. and Garapin, A.-C. (1979) Cloning of the active thymidine kinase gene of herpes simplex virus type I in E. coli K-12. Proc. Natl. Acad. Sci. USA, 76, p3755-3759.

Conway, 1. and Wickens, M.P. (1985) A sequence downstream of AAUAAA is required for efficient formation of SV40 late mRNA 3' termini in frog oocytes. Proc. Natl. Acad. Sci. USA, 82, p3949-3953.

Courey, A.J., Plon, S.E. and Wang, J.C. (1986) The use of psoralen-modified DNA to probe the mechanism of enhancer action. Cell, 45, p567-574.

Cregg, J.M., Tschopp, J.F. and Stillman, C. *et al.* (1987) High level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast, *Pichia pastoris*. Biotechnology, 5, p479-485.

Crouse, G.F., McEwan, R.N. and Pearson, M.L. (1983) Expression and amplification of engineered mouse DHFR minigenes. Mol. Cell Biol., 3, p257-266.

Cullen, B.R., Lomedico, P.T. and Ju, G. (1984) Transcriptional interference in avian retroviruses - implications for the promoter insertion mode of leukaemogenesis. Nature, 307, p241-245.

Cullen, B.R. (1987) Use of eucaryotic expression technology in the functional analysis of cloned genes. In 'Methods in Enzymology (a guide to molecular cloning techniques)' (eds. Berger, S.L. and Kimmel, A.R.), p684-704, Academic Press, N.Y., USA..

Dalgleish, R., (1987) Southern blotting. In 'Gene Cloning and Analysis' (ed. G.J. Boulnois), p45-60, Blackwell Scientific Publications, Oxford, England.

Danner, D. and Leder, P. (1985) Role of an RNA cleavage / poly(A) addition site in the production of membrane-bound and secreted IgM mRNA. Proc. Natl. Acad. Sci. USA, 82, p8658-8662.

Darnell Jr., J.E. (1982) Variety in the level of gene control in eucaryotic cells. Nature, 297, p365-371.

de Saint Vincent, R., Delbrück, B., Eckhart, S., Meinkoth, J., Vitto, L. and Wahl, G. (1981). The cloning and reintroduction into animal cells of a functional CAD gene, a dominant amplifiable genetic marker. Cell, 27, p267-278.

de Villiers, J., Olson, L., Banerji, J. and Schaffner, W. (1982) Analysis of the transcriptional effect. Cold Spring Harbor Symp. Quant. Biol., 47, p911-919.

Dick, J.E., Mägli, M.C., Phillips, R.A. and Bernstein, A. (1986) Genetic manipulation of haematopoietic stem cells with retrovirus vectors. Trends in Genet., 2, p165-170.

Doerfler, W. (1983) DNA methylation and gene activity. Ann. Rev. Biochem., 52, p93-124.

Doi, R.H., Wong, S.L. and Kawamura, F. (1986) Potential use of *Bacillus subtilis* for secretion and production of foreign proteins. Trends in Biotech., 4, p232-235.

Dougherty, J.P. and Temin, H.M. (1986) High mutation rate of a spleen necrosis virus-based retrovirus vector. Mol. Cell Biol., 7, p4387-4395.

Doyen. N., Leblond-Francillard, M., Holm, I., Dreyfus, M. and Rougeon, F. (1986) Analysis of promoter and enhancer cell type specificities and the regulation of Ig gene expression. Gene, 50, p321-331.

Druege, P.M., Klein-Hitpass, L., Green, S., Stack, G., Chambon, P. and Ryffel, G.U. (1986) Induction of estrogen responsiveness into mammalian cell lines. Nucl. Acids Res., 14, p9329-9337.

Dynan, W.S. (1987) What mechanisms underlie tissue-specific gene transcription? Trends in Genet., 3, p121-122.

Dzierzak, E.A., Papayannopoulou, T. and Mulligan, R.C. (1988) Lineage-specific expression of a human β-globin gene in murine bone marrow transplant recipients reconstituted with retrovirus-transduced stem cells. Nature, 331, p35-41.

Eagle, H. (1965) Metabolic controls in cultured mammalian cells. Science, 148, p42-51.

Early, D.H., Huang, H., Calame, K. and Hood, L. (1980) An Ig molecule heavy chain variable region is generated from 3 segments of DNA: V_H, D and J_H. Cell, 19, p981-992.

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L. (1980) Two mRNAs can be produced from a single immunoglobulin μ gene by alternative RNA processing pathways. Cell, 20, p313-319.

Echols, H. (1986) Multiple DNA-protein interactions governing high precision DNA transactions. Science, 233, p1050-1056.

Eckhart, L.A. and Birstein, B. (1985) Independent Ig class-switch events occurring in a single myeloma cell. Mol. Cell Biol., 5, p856-868.

Edery, I. and Sonenberg, N. (1985) Cap dependent RNA splicing in a HeLa nuclear extract. Proc. Natl. Acad. Sci. USA., 82, p7590-7594.

Edge, M.D. and Camble, R. (1984) Interferon synthesis in microorganisms. In 'Biotechnol. Genet. Eng. Rev.' (Ed. Russell, G.E.), 2, p215-252, Intercept, Newcastle, England.

Edlund, T., Walker, M.D., Barr, P.J. and Rutter, W.J. (1985) Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5' flanking sequences. Science, 23, p912-916.

Edmonds, M., Vaughan Jr., M.H. and Nakazato H. (1971) Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly labelled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship. Proc. Natl. Acad. Sci. USA, 68, p1336-1340.

Emerman, M. and Temin, H.M. (1984) Genes with promoters in retrovirus vectors can be independently supressed by an epigenetic mechanism. Cell, 39, p459-467.

Ephrussi, A., Church, G.M, Tonegawa, S. and Gilbert, W. (1985) B lineage-specific interactions of an Ig enhancer with cellular factors *in vivo*. Science, 227, p134-140.

Evans, M., Bradley, A. and Robertson, E.J. (1986) In 'Genetic Manipulation of the Mammalian Ovum and Early Embryo'. Cold Spring Harbor, N.Y., USA.

Evans, M. (1988) The use of the EK system for studies on homologous recombination. Presented at Biochem. Soc. meeting, Nottingham, England.

Evans, R.M. and Hollenberg, S.M. (1988) Zinc fingers: gilt by association. Cell, 52, p1-3.

Fahey, R.C., Hunt, S.J. and Windham, G.C. (1977) Cysteine and cystine content of proteins - differences between intracellular and extracellular proteins. J. Mol. Evol., 10, p155-160.

Faiferman, I. and Pogo, A.O. (1975) Isolation of nuclear ribonucleoprotein network that contains heterogeneous RNA and is bound to nuclear-envelope. Biochemistry, 14, p3808-3816.

Falck-Pedersen, E., Logan, J., Schenk, T. and Darnell, J.E. Jr. (1985) Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse β-major globin terminator element. Cell, 40, p897-905.

Falkner, F.G. and Zachau, H.G. (1984) Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature, 310, p71-74.

Folger,K.R., Wong, E.A., Wahl, G. and Capecchi, M.R. (1982) Patterns of integration of DNA microinjected into mammalian cells; evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell Biol., 2, p1372-1387.

Fojo, A.T., Whang-Peng, J., Gottesman, M.M. and Pastan, I. (1985) Amplification of DNA sequences in human multidrug resistant KB carcinoma cells. Proc. Natl. Acad. Sci., 82, p7661-7665.

Foster, J., Stafford, J. and Queen, C. (1985) An Ig promoter displays cell-type specificity independently of the enhancer. Nature, 315, p423-425.

Freedman, R.B. (1988) Post-translational modifications in the biosynthesis of secretory proteins. Presented at Biochem. Soc. meeting, Nottingham, England.

Freedman, R.B., Blackburn, E.C., Bullied, N.J., Hawkins, H.C., Murant, S.J. and Paver, J.L. (1988) The role of protein disulphide-isomerase in the expression of native proteins. Presented at Biochem. Soc. meeting, Nottingham, England.

Friedman, D. and Imperiale, M.J. (1987) RNA 3' end formation in the control of gene expression. Ann. Rev. Genet., 21, p453-488.

Fraser, N.W., Sehgal, P.B. and Darnell, J.F. (1979) Multiple discrete sites for premature RNA chain termination late in adenovirus-2 infection; enhancement by 5,6-dichloro-1-β-ribofuranosyl-benzimidazole. Proc. Natl. Acad. Sci. USA, 76, p2571-2575.

Furie, B. and Furie, B.C. (1988) The molecular basis of blood coagulation. Cell, 53, p505-518.

Furuichi, Y., LaFiandra, A. and Shatkin, A.J. (1977) 5'-terminal structure and mRNA stability. Nature, 266, p235-239.

Galfrè, G. and Milstein, C. (1981) Preparation of monoclonal antibodies: strategies and procedures. In 'Methods in Enzymology' (eds. Langone, J.L. and Van Vunakis, H.), 73, p1-46, Academic Press, N.Y., USA.

Galli, G., Guise, J.W., McDevitt, M.A., Tucker, P.W. and Nevins, J.R. (1987) Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted mu-chain expression during B-cell development. Genes Dev., 1, p471-481.

Garabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984) Messenger RNA splicing *in vitro*: an excised intervening sequence and a potential intermediate. Cell, 37, p415-427.

Garabowski, P.J., Seiler, S.R. and Sharp, P.A. (1985) A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell, 42, p345-353.

Gaub, M.P. *et al.* (1985) In 'Eucaryotic Transcription' (eds. Gluzman, Y. *et al*), p123-131, Cold Spring Harbour Laboratory, N.Y., USA.

Garcia, J.V, Bich-Thuy, L.-T., Stafford, J. and Queen, C. (1986) Synergism between Ig enhancers and promoters. Nature, 322, p383-385.

Gebara, M.M., Drevon, C., Harcourt, S.A., Steingrimsdottir, H., James, M.R., Burke, J.F., Arlett, C.F. and Lehmann, A.R. (1987) Inactivation of a

transfected gene in human fibroblasts can occur by deletion, amplification, phenotypic switching or methylation. Mol. Cell Biol., 7, p1459-1464.

Gellert, M. (1981) DNA topoisomerases. Ann. Rev. Biochem., 50, p879-910.

Gerster, T., Picard, D. and Schaffner, W. (1986) During B-cell differentiation, enhancer and transcription rate of IgH chain genes are high before mRNA accumulation. Cell, 45, p45-52.

Gerster, T., Matthias, P., Thali, M., Jirieny, J. and Schaffner, W. (1987) Cell type-specificity elements of the IgH chain gene enhancer. EMBO J., 6, p1323-1330.

Gibson, R., Schlesinger, S. and Kornfeld, S. (1979) The nonglycosylated glycoprotein of vesicular stomatitis virus is temperature-sensitive and undergoes intracellular aggregation at elevated temperatures. J. Biol. Chem., 254, p3600-3607.

Gibson, T.J. (1984) Studies on the Epstein-Barr virus genome. Ph.D thesis, University of Cambridge, England.

Gilbert, W. (1978) Why genes in pieces? Nature, 271, p501.

Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell, 33, p717-728.

Gilmore, R., Walter, P. and Blobel, G.J. (1982) Protein translocation across the endoplasmic reticulum. II. Isolation and characterisation of the signal recognition particle receptor. J. Cell Biol., 254, p470-477.

Gluzman, Y. (1981) SV40-transformed simian cells support the replication of early SV40 mutants. Cell, 23, p175-182.

Goldberg, G.I., Vanin, E.F., Zrolka, A.M. and Blattner, F.R. (1981) Sequence of the gene for the constant region of the mu chain of a balb/c mouse immunoglobulin. Gene, 15, p33-42.

Gopal, T.V. (1985) Gene transfer method for transient gene expression, stable transformation, and cotransformation of suspension cell cultures. Mol. Cell Biol., 5, p1188-1190.

Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982a) Recombinant genomes which express chloramphenicol acetyl transferase in mammalian cells. Mol. Cell Biol., 2, p1044-1051.

Gorman, C.M., Merlio, G.T, Willingham, M.C., Pastan, I and Howard, B.H. (1982b) The Rous sarcoma virus LTR is a strong promoter when introduced into a variety of eucaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA, 79, p6777-6781.

Gordon, K., Lee, E., Vitale, J., Smith, E., Westphal, H. and Hennighausen, L. (1987) Production of human tissue plasminogen activators in transgenic mouse milk. Biotechnology, 5, p1183-1187.

Goto, M., Akai, K., Murakima, A. *et al.* (1988) Production of human erythropoietin in mammalian cells: host-cell dependency of the biological activity of the cloned glycoprotein. Biotechnology, 6, p67-71.

Graham, F.L. and Eb Van der, A.J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology, 52, p456-467.

Grantham, R. (1980) Workings of the genetic code. Trends in Biochem. Sci., 5, p327-331.

Graves, R.A., Pandey, N.B., Chodchoy, N. and Marzluff, W.F. (1987) Translation is required for regulation of histone mRNA degradation. Cell, 48, p615-626.

Green, P.J., Pines, O. and Inogue, M. (1986) The role of anti-sense RNA in gene regulation. Ann. Rev. Biochem., 55, p569-597.

Greenberg, M.E., Hermanowski, A.L. and Ziff, E.B. (1986) Effect of protein synthesis inhibitors on growth-factor activation of c-*fos*, c-*myc* and actin gene-transcription. Mol. Cell Biol., 6, p1050-1057.

Gregor, P.D., Kobrun, B.J., Milcareck, C. and Morrison, S.L. (1986) Sequences 3' of IgH chain genes influence their expression. Immunological Rev., 89, p31-48.

Griffiths, B. (1986) Can cell culture medium costs be reduced? Strategies and possibilities. Trends in Biotech., 4, p268-272.

Gritz, L. and Davies, J. (1983) Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene, 25, p179-188.

Gros, P., Heriah, Y.B., Croop, J.M. and Housman, D.E. (1986) Isolation and expression of a cDNA that confers multidrug resistance. Nature, 323, p728-731.

Grosjean, H. and Fiers, W. (1982) Preferential codon usage in procaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene, 18, p199-202.

Gross, M.K., Kainz, M.S. and Merril, G.F. (1987) Introns are inconsequential to efficient formation of cellular thymidine kinase mRNA in mouse L cells. Mol. Cell Biol., 7, p4576-4581.

Gross, V., Steube, K., Tran-Thi, T.-A., Gerok, W. and Heinrich, P.C. (1988) The role of N-glycosylation for the clearance of rat liver secretory glycoproteins. Biochem. Soc. meeting, Nottingham, England.

Grosschedl, R. and Baltimore, D. (1985) Cell type specificity of Ig gene expression is regulated by at least three DNA sequence elements. Cell, 41, p885-897.

Grosveld, G.C., de Boer, E.C., Shewmaker, C.K. and Flavell, R.A. (1982) DNA sequences necessary for transcription of the rabbit β-globin gene *in vivo*. Nature, 295, p120-126.

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, p975-985.

Gruss, P. and Khoury, G. (1980) Rescue of a splicing defective mutant by insertion of a heterologous gene. Nature, 286, p634-637.

Gruss, P., Dhar, R. and Khoury, G. (1981) Simian virus 40 tandem repeat sequences as an element of the early promoter. Proc. Natl. Acad. Sci. USA, 78, p943-947.

Guarente, L. (1988) UASs and enhancers: common mechanisms of transcriptional activation in yeast and mammals. Cell, 52, p303-305.

Gunning, P., Leavitt, J., Muscat, G., Ng, S.-Y. and Kedes, L. (1987) A human β-actin expression vector system directs high level accumulation of anti-sense transcripts. Proc. Natl. Acad. Sci. USA, 84, p4381-4385.

Gwynne, D.I., Buxton, F.P., Sills, M. and Davies, R.W. (1988) Genetically engineered expression and secretion of foreign proteins from *Aspergillus nidulans*. Presented at Biochem. Soc. meeting, Nottingham, England.

Haas, I.G. and Wabl, M. (1983) Immunoglobulin heavy chain binding proteins. Nature, 306, p387-389.

Hamer, D.H. and Walling, M.J. (1982) Regulation *in vivo* of a cloned mammalian gene: cadmium induces the transcription of a mouse metallothionein gene in SV40 vectors. J. Mol. Appl. Genet., 1, p273-288.

Hanahan, D., Lane, D., Lipsich, L., Wigler, M. and Botcham, M. (1980) Characterisation of a SV40-plasmid recombinant and its movement into and out of the genome of a murine cell. Cell, 21, p127-139.

Hanahan, D., (1983) Studies on transformation of *E. coli* with plasmids. J. Mol. Biol. 166, p557-580.

Hanahan, D., (1985) Rapid and efficient transformation of *E. coli*. In 'DNA Cloning' (ed. Glover, D.M.), 1, p109-114, IRL Press; Oxford, England and Washington, D.C., USA.

Hanamada, H. (1986) Activation of an enhancerless gene by chromosomal integration. Mol. Cell Biol., 6, p4179-4184.

Harris, T.J.R. (1987) Second-generation plasminogen activators. Protein Eng., 1, p449-458.

Hartung, S., Jaenisch, R. and Breindl, M. (1986) Retrovirus insertion inactivates mouse α1(I) collagen gene by blocking initiation of transcription. Nature, 320, p365-367.

Hashimoto, C. and Steitz, J.A. (1986) A small nuclear ribonucleoprotein associated with the AAUAAA polyadenylation signal *in vitro*. Cell, 45, p315-318.

Hausler, B. and Somerville, R.L. (1979) Interaction *in vivo* between strong closely spaced constitutive promoters. J. Mol. Biol., 127, p353-356.

Hayward, B.E., Hussain, A., Wilson, R.H., Woodcock, V., McIntosh, B. and Harris, T.J.R. (1986) The cloning and nucleotide sequence of cDNA for an amplified glutamine synthetase gene from the chinese hamster. Nucl. Acids Res., 14, p999-1008.

Heijne von, G. (1983) Pattern of amino acids near signal-sequence cleavage sites. Eur. J. Biochem., 133, p17-21.

Hendricks, M.R. and Banker, M.J. (1987) Modulation of expression of transfected immunoglobulin genes by DNA in the 5' untranslated region: a relationship between RNA accumulation and the size of the 5'-proximal exon. J. Cell Biochem. suppl., 11C, p125.

Hendricks, M.B., Banker, M.J. and McLaughlin, M. (1988) A high efficiency vector for expression of foreign genes in myeloma cells. Gene, 64, p43-51.

Heuvel van, M., Bosveld, I.J., Luyten, W., Trapman, J. and Zwartoff, E.C. (1986) Transient expression of murine interferon alpha genes in mouse and monkey cells. Gene, 45, p159-165.

Holland, I.B. (1988) Mechanism of secretion of *E. coli* haemolysin. Presented at Biochem. Soc. meeting, Nottingham, England.

Holloman, W., Wiegand, R., Hoessli, C. and Radding, C. (1975) Uptake of homologous single-stranded fragments by superhelical DNA: a possible mechanism for initiation of genetic recombination. Proc. Natl. Acad. Sci. USA, 72, p2394-2398.

Hope, I.A. and Struhl, K. (1986) Functional dissection of a eucaryotic transcriptional activator protein, GCN4, in yeast. Cell, 46, p885-894.

Howard, B.H. (1983) Vectors for introducing genes into cells of higher eucaryotes. Trends in Biochem. Sci., 8, p209-212.

Hu, N. and Messing, J. (1982) The making of strand specific M13 probes. Gene, 17, p271-277.

Hunt, T. (1988) Controlling mRNA lifespan. Nature, 334, p567-568.

Hwang, S.L.-H. and Gilboa, E. (1984) Expression of genes introduced into cells by retroviral infection is more efficient than that of genes introduced into cells by DNA transfection. J. Virol., 50, p417-424.

Ikemura, T. (1981) Correlation between the abundance of *E. coli* transfer RNAs and the occurence of the respective codon in its protein genes. J. Mol. Biol., 146, p1-21.

Imler, J.-L., Lemaire, C., Wasylyk, C. and Wasylyk, B. (1987) Negative regulation contributes to tissue specificity of the IgH enhancer. Mol. Cell Biol., 7, p2558-2567.

Inada, Y., Yoshimoto, T., Matasushima, A. and Saito, Y. (1986) Engineering physicochemical and biological properties of proteins by chemical modification. Trends in Biotech., 4, p68-73.

Inlow, D., Harano, D. and Maiorella, B. (1987) Large scale insect cultures for recombinant protein production. Presented at Symposium on Strategies in Cell-Culture Scale Up. American Chemical Society national meeting, New Orleans, USA.

Ish-Horowicz, D. (1982) Rapid, small-scale isolation of plasmid DNA; alkaline lysis method. In 'Molecular Cloning - a laboratory manual' (eds. Maniatis, T., Fritsch, E.F. and Sambrook, J.), p368, Cold Spring Harbor Lab., N.Y., USA.

Jaenisch, R., Johner, D., Nobis, P., Simon, I., Lohler, J., Harkers, K. and Grotkopp, D. (1981) Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell, 24, p1541-1552.

Jackson, D.A. and Cook, P.R. (1985) A general method for preparing chromatin containing intact DNA. EMBO J., 4, p919-925.

Jackson, D.A. (1986) Organisation beyond the gene. Trends in Biochem. Sci., 11, p249-252.

Jang, S.K. *et al.* (1988) J. Virol. (in press).

Jeang, K.-T. and Khoury, G. (1988) The mechanistic role of enhancer elements in eucaryotic transcription. Bioessays, 8, p104-107.

Johnson, M.R., Norman, C., Reeve, M., Scully, J. and Proudfoot, N.J. (1986) Tripartite sequences within and 3' to the sea urchin H2A histone gene display properties associated with a transcriptional termination process. Mol. Cell Biol., 6, p4008-4018.

Jones, N.C. (1986) Negative regulation of enhancers. Nature, 321, p202-203.

Jongstra, J., Rendelhuber, T.L., Oudert, P., Benoist, C., Chae, C.B., Jeltsch, J.M., Mathis, D.T. and Chambon, P. (1984) Induction of altered chromatin structures by simian virus 40 enhancer and promoter elements. Nature, 307, p708-714.

Joyner, A., Keller, G., Phillips, R.A. and Bernstein, A. (1983) Retrovirus-mediated transfer of a bacterial gene into mouse haematopoietic progenitor cells. Nature, 305, p556-558.

Jung, A., Sippel, A.E., Grez, M. and Schutz, G. (1980) Exons encode functional and structural units of chicken lysozyme. Proc. Natl. Acad. Sci. USA, 77, p5759-5763.

Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M. and Perry, H. (1983) Sequences of proteins of Immunological Interest. Health and Human Services Department, NIH, Bethesda, USA.

Kadesch, T. and Berg, P. (1986) Effects of the position of the SV40 enhancer on expression of multiple transcription units in a single plasmid. Mol. Cell Biol., 6, p2593-2601.

Kadesch, T., Zevos, P. and Ruezinsky, D. (1986) Functional analysis of the murine IgH enhancer; evidence for negative control of cell-type specificity. Nucl. Acids Res., 14, p8209-8221.

Kaetzel, D.M., Browne, J.K., Wondisford, F., Nett, T.M., Thomason, A.R. and Nilson, J.H. (1985) Expression of biologically active bovine lutenising hormone in chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA, 82, p7280-7283.

Kaufman, R.J. and Sharp, P.A. (1982) Amplification and expression of sequences cotransfected with modular dihydrofolate reductase cDNA gene. J. Mol. Biol., 159, p601-621.

Kaufman, R.J., Wasley, L.C., Spiliotes, A.J., Gossels, S.D., Latt, S.A., Larsen, G.R. and Kay, R.M. (1985a) Coamplification and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in chinese hamster ovary cells. Mol. Cell Biol., 5, p1349-1357.

Kaufman, R.J. (1985b) Identification of the components necessary for adenovirus translational control and their utilisation in cDNA expression vectors. Proc. Natl. Acad. Sci. USA, 82, p689-693.

Kaufman, R.J, Mutrtha, P., Ingolia, D.E., Yeung, C. and Kellems, R.E. (1986) Selection and coamplification of heterologous genes encoding adenosine deaminase in mammalian cells. Proc. Natl. Acad. Sci. USA, 83, p3136-3140.

Kelley, D.E. and Perry, R.P. (1986) Transcriptional and post-transcriptional control of Ig mRNA production during B lymphocyte development. Nucl. Acids Res., 14, p5431-5437.

Kelly, R.B. (1985) Pathways of protein secretion in eucaryotes. Science, 230, p25-32.

Kim, S.K. and Wold, B.J. (1985) Stable reduction of thymidine kinase activity in cells expressing high levels of anti-sense RNA. Cell, 42, p729-736.

Kingsman, S.M., Kingsman, A.J. and Mellor, J. (1987) The production of mammalian proteins in *Saccharomyces cerevisiae*. Trends in Biotech., 5, p53-56.

Kinosita, K. and Tsong, T.Y. (1977) Voltage-induced conductance in human erythrocyte membranes. Biochem. Biophys. Acta, 471, p227-242.

Kioussis, D. (1988) Integration-site-dependent expression of the human CD2 gene in transgenic mice. Presented at Biochem. Soc. meeting, Nottingham, England.

Klein, S., Gerster, T., Picard, D., Radbruch, A. and Schaffner, W. (1985) Evidence for transient requirement of the IgH enhancer. Nucl. Acids Res., 13, p8901-8911.

Köhler, G. and Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. Nature, 256, p495-497.

Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984) Recognition of cap structure in splicing *in vitro* of mRNA precursors. Cell, 38, p731-736.

Kornfield, R. and Kornfield, S. (1985) Assembly of asparagine-linked oligosaccharides. Ann. Rev. Biochem., 259, p672-686.

Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. Microbial Rev., 47, p1-45.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eucaryotic ribosomes. Cell, 44, p283-292.

Krainer, A.R., Maniatis, T., Ruskin, B. and Green, M.R. (1984) Normal and mutant human β-globin pre-mRNAs are faithfully and efficiently spliced *in vitro*. Cell, 36, p993-1005.

Krieg, P., Strachen, R., Wallis, E., Tabe, L. and Colman, A. (1984) Efficient expression of cloned and complementary DNAs for secretory proteins after injection into *Xenopus* oocytes. J. Mol. Biol., 180, p615-643.

Kurland, C.G. (1987) Strategies for efficiency and accuracy in gene expression. Trends in Biochem. Sci., 12, p126-128.

Land, H., Grez, M., Hauser, H., Lindenmaier, W. and Schutz, G. (1981) 5'-terminal sequences of eucaryotic mRNA can be cloned with high efficiency. Nucl. Acids Res., 9, p2251-2266.

Laemmli, U.K., (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. Nature, 227, p680-685.

Laimins, L., Holmgren-König, M. and Khoury, G. (1986) Transcriptional 'silencer' element in rat repetitive sequences associated with the rat insulin 1 gene locus. Proc. Natl. Acad. Sci. USA, 83, p3151-3155.

Lanoix, J. and Acheson, N.H. (1988) A rabbit β-globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA. EMBO J., 7, p2515-2522.

Laskov, R. and Scharff, M.D. (1970) Synthesis, assembly and secretion of gamma globulin by mouse myeloma cells. I. Adaption of the MPC-11 tumour to culture cloning and characterisation of gamma globulin subunits. J. Exp. Med., 131, p515-541.

Lasky, M. and Botchan, M.R. (1984) Characteristics of the bovine papillomavirus plasmid maintenance sequences. Cell, 36, p391-401.

Lee, F., Mulligan, R., Berg, P. and Ringold, G. (1981) Glucocorticoids regulate expression of dihydrofolate reductase cDNA im mouse mammary tumour virus chimaeric plasmids. Nature, 294, p228-232.

Legerski, R.J., Hodnett J.L. and Gray Jr. H.B. (1978) Extracellular nucleases of Pseudomonas Bal 31 III. Use of the double-strand deoxyribonuclease activity as the basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. Nucl. Acids Res., 5, p1445-1464.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steiz, J.A. (1980) Are snRNPs involved in splicing? Nature, 283, p220-224.

Levinson, B., Khoury, G., Vande Woude, G. and Gruss, P. (1982) Activation of SV40 genome by 72-base pair tandem repeats of a Moloney sarcoma virus. Nature, 295, p568-572.

Lewis, J.A. (1986) Structure and expression of the chinese hamster thymidine kinase gene. Mol. Cell Biol., 6, p1998-2010.

Liang, S.M., Allet, B., Rose, K., Hirschi, M., Liang, C.-M. and Thatcher, D.K. (1985) Characterisation of human interleukin 2 derived from *E. coli*. Biochem. J., 229, p429-439.

Lichtenstein, C. (1988) Anti-sense RNA as a tool to study plant gene expression. Nature, 333, p801-802.

Logan, J. and Shenk, T. (1984) Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. Proc. Natl. Acad. Sci. USA, 81, p3655-3659.

Lord, S.J. (1985) Expression of a cloned human fibrinogen cDNA in *E. coli* - synthesis of an A-alpha-polypeptide. DNA, 4, p33-38.

Lowry, P.J. and Chadwick, A. (1970) Interrelation of some pituitary hormones. Nature, 226, p219-224.

Lucklow, V.A. and Summers, M.D. (1988) Trends in the development of baculovirus expression vectors. Biotechnology, 6, p47-55.

Lundblad, A., Steller, R., Kabat, E.A., Hirst, J.W., Weigert, M.G. and Cohn, M. (1972) Immunochemical studies on mouse myeloma proteins with specificity for dextran or for levan. Immunochemistry, 9, p535-544.

Lusky, M. and Botchan, M.R. (1984) Characterisation of the bovine papillomavirus plasmid maintenance sequences. Cell, 36, p391-401.

MacGregor, G. and Burke, J.F. (1987) Stability of a bacterial gene in a Bovine Papilloma Virus based shuttle vector maintained extrachromosomally in mammalian cells. J. Gen. Virol., 67, p247-252.

Mains, P.E. and Sibley, C.H. (1983) The requirement of light chain for the surface deposition of the heavy chain immunoglobulin M. J. Biol. Chem., 258, p5027-5033.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (eds.) (1982) Molecular Cloning - a laboratory manual. Cold Spring Harbor Labotatory, N.Y., USA.

Marzluff, W.F. and Pandey, N.B. (1988) Multiple regulatory steps control histone mRNA concentrations. Trends in Biochem., 13, p49-52.

Mann, R., Mulligan, R.C. and Baltimore, D. (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell, 33, p153-159.

Mason, J.O., Williams, G.T. and Neuberger, M.S. (1985) Transcription cell type specificity is conferred by an Ig V_H gene promoter that includes a functional consensus sequence. Cell, 41, p479-487.

Maxam, A. and Gilbert, W. (1980) Sequencing end-labelled DNA with base specific chemical cleavages. In 'Methods in Enzymology' (eds. Grossman, L. and Moldave, K.), 65, p499-560, Academic Press, N.Y., USA.

Mayo, K.E., Warren, R. and Palmiter, R.D. (1982) The mouse metallothionein-I gene is transcriptionally regulated by cadmium following transfection into human or mouse cells. Cell, 29, p99-108.

Margolskee, R.F., Kavathas, P. and Berg, P. (1988) Epstein-Barr virus shuttle vector for stable episomal replication of cDNA expression libraries in human cells. Mol. Cell Biol., 8, p2837-2847.

Marston, F.A.O. (1986) The purification of eucaryotic polypeptides synthesised in *E. coli*. Biochem. J., 240, p1-12.

Marzluff, W.F. and Pandey, N.B. (1988) Multiple regulatory steps control histone mRNA concentrations. Trends in Biochem. Sci., 13, p49-52.

Mather, E.L., Nelson, K.J., Haimovich, J. and Perry, R.P. (1984) Mode of regulation of Ig μ and δ-chain expression varies during B-lymphocyte maturation. Cell, 36, p329-338.

McDonell M.W., Simon M.N. and Studier F.W. (1977) Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol., 110, p119-146.

McIvor, R.S., Goddard, J.M., Simonsen, C.C. and Martin, D.W. (1985) Expression of a cDNA sequence encoding human purine nucleoside phosphorylase in rodent and human cells. Mol. Cell Biol., 5, p1349-1357.

McKnight, S.L., Gavis, E.R. and Kingsbury, R. (1981) Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell, 25, p385-398.

McKnight, S.L., Kingsbury, R.C., Spence, A. and Smith, M. (1984) The distal transcriptional signals of the herpes virus tk gene share a common hexanucleotide control sequence. Cell, 37, p253-262.

McKnight, S.L. and Tjian, R. (1986) Transcriptional selectivity of viral genes in mammalian cells. Cell, 46, p795-805.

McLauchlan, J., Gaffrey, D., Whitton, J.L. and Clements, B. (1985) The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. Nucl. Acids Res., 13, p1347-1365.

McReynolds, L.A., Catterall, J.F. and O'Malley, B.W. (1977) The ovalbumin gene: cloning of a complete ds-cDNA in a bacterial plasmid. Gene, 2, p217-231.

Meck, R.L., Walsch, R.D. and Palmiter, J. (1982) The signal sequence of ovalbumin is located near the NH_2 terminus. J. Biol. Chem., 257, p12245-12251.

Melchers, F. and Andersen, J. (1984) B cell activation: three steps and their variations. Cell, 37, p715-720.

Mendelson, E., Landsman, D., Druckmann, S. and Bustin, M. (1986) Immunofractionation of chromatin regions associated with histone H1°. Eur. J. Biochem., 160, p253-260.

Merten, O.W. (1987) Concentrating mammalian cells I. Large-scale animal cell culture. Trends in Biotech., 5, p230-237.

Meyer, D.I., Krause, E. and Doberstein, B. (1982) Secretory protein translocation across membranes - the role of the 'docking protein'. Nature, 297, p647-650.

Milcareck, C. and Hall, B. (1985) Cell specific expression of secreted versus membrane form of Igγ2B mRNA involves selective use of alternate polyadenylation sites. Mol. Cell Biol., 3, p1317-1332.

Miller, A.D., Ong, E.S., Rosenfeld, M.G., Verma, I. and Evans R.M. (1984) Infectious and selectable retrovirus containing an inducible rat growth hormone minigene. Science. 225, p993-998.

Miller, T.J., Stephens, D.L. and Mertz, J.E. (1982) Kinetics of accumulation and processing of SV40 RNA accumulation in *Xenopus laevis* injected with SV40 DNA. Mol. Cell Biol., 2, p1581-1594.

Minie, M.E. and Koshland, M.E. (1986) Accessibility of the promoter in the J chain is regulated by chromatin changes during B cell differentiation. Mol. Cell Biol., 6, p4031-4038.

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, p223-232.

Montgomery, D.L., Hall, B.D., Gillam, S. and Smith, M. (1978) Identification and isolation of yeast cytochrome c gene. Cell, 14, p673-680.

Morris, T., Marashi, F., Weber, L., Hickey, E., Greenspan, D., Bonner, J., Stein, J. and Stein, G. (1986) Involvement of the 5' leader sequence in coupling the stability of a human H-3 histone mRNA with DNA replication. Proc. Natl. Acad. Sci. USA. 83, p981-985.

Mount, S.M., Petterson. I., Hinterberger, M., Karmes, A. and Steiz, J.A. (1983) The U1 small nuclear RNA-protein complex selectively bind a 5' splice site *in vitro*. Cell, 33, p509-518.

Moye, W.S. and Zalakin, H. (1987) Role of the TGACT repeat sequence in yeast-TRP5 promoter. J. Biol. Chem., 262, p3069-3614.

Mulligan, R.C. and Berg, P. (1980) Expression of a bacterial gene in mammalian cells. Science, 209, p1422-1427.

Mulligan, R.C. and Berg, P., (1981a) Selection for animal cells that express the *E. coli* gene coding for xanthine-guanine phosphoribosyltransferase. Proc. Natl. Acad. Sci. USA, 78, p2072-2076.

Mulligan, R.C. and Berg, P. (1981b) Factors governing the expression of a bacterial gene in mammalian cells. Mol. Cell Biol., 1, p449-459.

Mulligan, R.C. (1988) Retrovirus-mediated gene transfer and expression in somatic cells. Presented at Biochem. Soc. meeting, Nottingham, England.

Müllner, E.W. and Kühn, L.C. (1988) A stem-loop in the 3' untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. Cell, 53, p815-825.

Murray, M.J., Kaufman, R.J., Latt, S.A. and Weinberg, R.A. (1983) Construction and use of the dominant selectable marker: a Harvey Sarcoma virus-dihydrofolate reductase chimera. Mol. Cell Biol., 3, p32-43.

Nagahari, K., Kanaya, S., Munkata, K., Aoyagi, V. and Mizushima, S. (1985) Secretion into the culture-medium of a foreign gene product from *E. coli*; use of the *ompF* gene for secretion of human beta-endorphin. EMBO J., 4, p3589-3592.

Narita, K. (1975) In 'Protein Structure and Function' (eds. Fanatsu, M., Hiromi, K., Imahori, K., Murachi, K. and Narita, K.), 2, p227-259. Wiley, N.Y., USA.

Nelsestuen, G.L., Zytkovicz, T.H. and Howard, J.B. (1974) The mode of action of vitamin K. Identification of γ-carboxyglutamic acid as a component of prothrombin. J. Biol. Chem., 249, p5648-5656.

Nelson, K.J., Haimovich, J. and Perry, R.P. (1983) Characterisation of productive and sterile transcripts from the IgH chain locus: Processing of μ_m and μ_s mRNAs. Mol. Cell Biol., 3, p1317-1332.

Neuberger, M.S. (1983) Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. EMBO J., 2, p1373-1378.

Neuberger, M.S. (1985) Making novel antibodies by expressing transfected Ig genes. Trends in Biochem. Sci., 10, p347-349.

Neurath, H. and Walsh, K.A. (1976) Role of proteolytic enzymes in biological regulation. Proc. Natl. Acad. Sci. USA., 73, p3825-3828.

Nordheim, A. and Rich, A. (1983) Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature, 303, p674-679.

Nossal, G.J.V. and Pike, B.L. (1972) Differentiation of B lymphocyte from stem cell precursors. Adv. Exp. Med. Biol., 29, p11-20.

Novick, P., Field, C. and Schekman, R. (1980) Identification of 23 complementation groups required for posttranslational events in the yeast secretory pathway. Cell, 21, p205-215.

Oberto, J. and Davison, D. (1985) Expression of chicken egg-white lysozyme by *Saccharomyces cerevisiae*. Gene, 40, p57-65.

Oi, V.T., Morrison, S.L., Herzenberg, L.A. and Berg, P. (1983) Immunoglobulin gene expression in transfected lymphoid cells. Proc. Natl. Acad. Sci. USA, 80, p825-829.

Okayama, H. and Berg, P. (1983) A cDNA cloning vector that permits expression of cDNA inserts into mammalian cells. Mol. Cell Biol., 3, p280-289.

Olson, K.C., Fenno, J., Lin, N., Harkins, R.N., Snider, C., Kohr, W.H., Ross, M.J., Fodge, D., Prender, N. and Stebbing, N. (1981) Purified human growth hormone from *E. coli* is biologically active. Nature, 293, p408-411.

Orci, L., Halban, P., Amherdt, M., Ravazzola, M., Vassalli, J.-D. and Perrelet, A. (1984) A clathrin coated, golgi-related compartment of the insulin secreting cell accumulates proinsulin in the presence of monensin. Cell, 39, p39-47.

Orci, L., Glick, B.S. and Rothman, J.E. (1986) A new type of coated vesicular carrier that appears not to contain clathrin: its possible role in protein transport within the golgi stack. Cell, 46, p171-184.

Padgett, R.A., Grabowski, P.J., Konarska, M.M. and Sharp, P.A. (1985) Splicing messenger RNA precursors: branch sites and lariat RNAs. Trends in Biochem. Sci., 10, p154-157.

Palade, G. (1975) Intracellular aspects of protein synthesis. Science, 189, p347-358.

Palatnik, C.M., Wilkins, C. and Jacobson, A. (1984) Translational control during early Dictyostelium development: possible involvement of poly(A) sequences. Cell, 36, p1017-1025.

Palmiter, R.D., Gagnon, J., Evicsson, L.H. and Walsh, K.A. (1977) Precursor of egg white lysozyme. Amino acid sequence of an NH_2-terminal extension. J. Biol. Chem., 252, p6386-6393.

Pavirani, A., Meulien, P., Harrer, H., Schamber, F., Dott, K., Villeval, D., Cordier, Y., Wiesel, M.L., Marzurier, C., Van de Pol, H., Piquet, Y., Cazenave, J.P. and Lecoq, J.P. (1987) Choosing a host cell for active recombinant factor VIII production using vaccinia virus. Biotechnology, 5, 389-391.

Pavlakis, G.N. and Hamer, D.H. (1983) Regulation of metallothionein growth hormone hybrid gene in bovine papillomavirus. Proc. Natl. Acad. Sci. USA, 80, p1527-1531.

Payvar, F., DeFranco, D., Firestone, G.L., Edgar, B., Wrange, O., Okret, S., Gustaffson, J.A. and Yamamoto, K.R. (1983) Sequence specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. Cell, 35, p381-392.

Peabody, D.S. and Berg, P. (1986) Termination-reinitiation occurs in the translation of mammalian cell mRNA. Mol. Cell Biol., 6, p2695-2703.

Pelletier, J. and Sonenberg, N. (1988) Internal initiation of translation of eucaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature, 324, p320-325.

Pelletier, J. and Sonenburg, N. (1985) Insertion mutagenesis to increase secondary structure within the 5' noncoding region of eucaryotic mRNAs reduces translational efficiency. Cell, 40, p515-526.

Perkins, K.K., Furneaux, H.M. and Hurwitz, J. (1986) RNA splicing products formed with isolated fractions from HeLa cells are associated with fast sedimenting complexes. Proc. Natl. Acad. Sci. USA, 83, p887-891.

Perkus, M.E., Piccini, A., Lipinskas, B.R. amd Paoletti, E. (1985) Recombinant vaccinia virus: immunisation against multiple pathogens. Science, 229, p981-984.

Perlman, D. and Halvorson, O.H. (1983) A putative signal peptidase recognition site and sequence in eucaryotic and procaryotic signal peptides. J. Mol. Biol., 167, p391-409.

Perry, R.P. and Kelley, D.E. (1979) Ig mRNA in murine cell lines that have characteristics of immature B lymphocytes. Cell, 18, p1333-1339.

Perucho, M., Hanahan, D. and Wigler, M. (1980) Genetic and physical linkage of exogenous sequences in transformed cells. Cell, 22, p309-317.

Peterson, M.L. and Perry, R.P. (1986) Regulated production of μ_m and μ_s mRNA requires linkage of the poly(A) addition sites and is dependent on the lengths of the μ_s-μ_m intron. Proc. Natl. Acad. Sci. USA, 83, p8883-8887.

Pfeffer, S.R. and Rothman, J.E. (1987) Biosynthetic protein transport and sorting by the endoplasmic reticulum and golgi. Ann. Rev. Biochem., 56, p828-852.

Plon, S.E. and Wang, J.C. (1986) Transcription of the human B-globin gene is stimulated by an SV40 enhancer to which it is physically linked but topologically uncoupled. Cell, 45, p575-580.

Potter, H., Weir, L. and Leder, P. (1984) Enhancer-dependent expression of human κ immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. Proc. Natl. Acad. Sci. USA, 81, p7161-7165.

Pratt, D and Subramani, S. (1983) Nucleotide sequence of the *E. coli* xanthine-guanine phosphoribosyl transferase gene. Nucl. Acids Res., 11, p8817-8823.

Proudfoot, N.J. (1986) Transcriptional interference and termination between duplicated α-globin gene constructs suggests a novel mechanism for gene regulation. Nature, 322, p562-565.

Proudfoot, N.J. and Brownlee, G.G. (1976) 3' non-coding region sequences in eucaryotic mRNA. Nature, 263, p211-214.

Pruss, G.J. and Drlica, K. (1986) Topoisomerase I mutants; the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling. Proc. Natl. Acad. Sci., 83, p8952-8956.

Ptashne, M. (1986) Gene regulation by proteins acting nearby and at distance. Nature, 322, p697-701.

Purvis, I., Bettany, A.J.E., Santiago, T.C. Coggins, J.R., Duncan, K., Eason, R. and Brown, A.J.P. (1987) The efficiency of folding of some proteins is increased by controlled rates of translation *in vivo*. J. Mol. Biol., 193, p413-417.

Queen, C. and Baltimore, D. (1983) Immunoglobulin gene transcription is activated by downstream sequence elements. Cell, 33, p741-748.

Rabbits, P.H., Forster, A., Stinson, M.A. and Rabbits, T.H. (1985) Truncation of exon 1 from the c-*myc* gene results in prolonged c-*myc* stability. EMBO J., 4, p3727-3733.

Ramabhadran, T.V. (1987) Products from genetically engineered mammalian cells: benefits and risk factors. Trends in Biotech., 5, p175-179.

Rassoulzadegan, M., Binetray, B. and Cuzin, F. (1982) The development of spheroplast fusion as a transfection procedure for eucaryotic cells. Nature, 295, p257-259.

Raghow, R. (1987) Regulation of messenger RNA turnover in eucaryotes. Trends in Biochem. Sci., 12, p358-360.

Ratafia, M. (1987) Mammalian cell culture: world wide activities and markets. Biotechnology, 5, p692-694.

Reddy, V.B., Ghosh, P.K., Lebowitz, P., Piatak, M. and Weissman, S.M. (1979) Simian virus 40 early mRNAs - 1. Genomic localisation of 3' and 5' termini of two major splices in mRNA from transformed and lytically infected cells. J. Virology., 30, p279-296.

Reeves, R. (1984) Transcriptionally active chromatin. Biochem. Biophys. Acta, 782, p343-393.

Renaut, E., Stanssens, P. and Fiers, W. (1981) Plasmid vectors for high efficiency expression controlled by the p_L promoter of coliphage lambda. Gene, 15, p81-93.

Reuther, J.E., Maderious, A., Lavery, D., Logan, J., Manfu, S. and Chen-Kiang, S. (1985) Cell-specific expression of secreted versus membrane forms of Igγ2b mRNA involves selective use of alternate polyadenylation sites. Mol. Cell Biol., 5, p2514-2250.

Reuveny, S., Velez, D., Miller, L. and Macmillan, J.D. (1986) Factors affecting cell growth and monoclonal antibody production in stirred reactors. J. Immunol. Methods, 86, p61-69.

Rio, D.C., Clark, S.G. and Tjian, R. (1985) A mammalian host-vector system that regulates expression and amplification of transfected genes by temperature induction. Science, 227, p23-28.

Richter, D. (1983) Vasopressin and oxytocin are expressed as polyproteins. Trends in Biochem. Sci., 8, p278-281.

Roberts, J.M. and Axel, R. (1982) In 'Gene Amplification' (ed. Schimke, R.). Cold Spring Harbor Laboratory, N.Y., USA.

Robertson, E.J. (1986) Pluripotential stem cell lines as a route into the mouse germ line. Trends in Genet., 2, p9-13.

Robins, D.M., Ripley, S. Henderson, A.S. and Axel,R. (1981) Transforming DNA integrates into the host chromosome. Cell, 23, p29-39.

Ross, J., Kobs, G., Brewer, G. and Peltz, S.W. (1987) Properties of the exonuclease activity that degrades H4 histone mRNA. J. Biol. Chem., 262, p9374-9381.

Roth, R.A., Mather, E.L. and Koshland, M.E. (1979) Intracellular events in the differentiation of B lymphocytes to pentamer IgM synthesis. In 'Cells of Immunoglobulin Synthesis' (Eds. Vogel, H. and Pernis, B.), p141-151, Academic Press, N.Y., USA.

Ryoji, M. and Worcel, A. (1984) Chromatin assembly in *Xenopus* oocytes: *in vivo* studies. Cell, 37, p1227-1232.

Sanger, F. Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA, 74, p5463-5467.

Sairam, M.R. and Bhargavi, G.N. (1985) A role for glycosylation of the α subunit in transduction of a biological signal in glycoprotein hormones. Science, 229, p65-67.

Sambucetti, L.C., Schaber, M., Kramer, R., Crowl, R. and Curran, T. (1986) The *fos* gene product undergoes extensive post-translational modification in eucaryotic but not in procaryotic cells. Gene, 43, p69-77.

Saragoti, S., Mayne, G. and Yaniv, M. (1980) Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding the late RNA. Cell, 20, p65-73.

Sarver, N., Gruss, P., Law, M.-F., Khoury, G. and Howley, P.M. (1981) Bovine papillomavirus deoxyribonucleic acid: a novel eucaryotic cloning vector. Mol. Cell. Biol., 1, p486-496.

Sassone-Corsi, P. and Borrelli, E. (1986) Transcriptional regulation by *trans*-acting factors. Trends in Genet., 11, p215-219.

Scangos, G. and Ruddle, F.H. (1981) Mechanisms and applications of DNA mediated gene transfer in mammalian cells - a review. Gene, 14, p1-10.

Schaffner, W. (1980) Direct transfer of cloned genes from bacteria to mammalian cells. Proc. Natl. Acad. Sci. USA, 77, p2163-2167.

Scharff, M.D., Bargellesi, A., Baumal, R., Buxbaum, J., Coffino, P. and Laskov, R. (1970) Variations in the synthesis and assembly of an immunoglobulin by mouse myeloma cells: a genetic and biochemical analysis. J. Cell Physiol., 76, p331-348.

Scheiness, D., Puckett, L. and Darnell, J.E. (1975) Possible relationship of poly(A) shortening to mRNA turnover. Proc. Natl. Acad. Sci. USA, 72, p1077-1081.

Schneider, R.J., Weinberger, C. and Shenk, T. (1984) Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. Cell, 37, p291-298.

Scholkat, U., Bohman, D., Schöler, H. and Gruss, P. (1986) Nuclear factors binding specific sequences within the Ig enhancer interact differentially with other elements. EMBO J., 5, P3251-3258.

Schröder, H.C., Zahn, R.K. and Müller, W.E.G. (1982) Role of actin and tubulin in the regulation of poly(A) polymerase endoribonuclease IV complex from calf thymus. J. Biol. Chem., 257, p2305-2309.

Schröder, H.C., Rottmann, M., Bachmann, M. and Müller, W.E.G. (1986) Purification and characterisation of the major nucleoside triphosphate from rat liver nuclear envelope. J. Biol. Chem., 261, p663-668.

Schröder, H.C., Trölltsch, D., Friese, U., Bachmann, M. and Müller, W.E.G. (1987) Mature mRNA is selectively released from the nuclear matrix by an ATP/dATP-dependent mechanism sensitive to topoisomerase inhibitors. J. Biol. Chem., 262, p8917-8925.

Sefton, B.M. and Buss, J.E. (1987) The covalent modification of eucaryotic proteins with lipid. J. Cell Biol., 104, 1449-1453.

Serfling, E., Jasin, M. and Schaffner, W. (1985) Enhancers and eucaryotic gene transcription. Trends in Genet., 10, p224-230.

Shatkin, A.J. (1976) Capping of eucaryotic mRNAs. Cell, 9, p645-653.

Shaw and Kamen, R. (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates mRNA selective degradation. Cell, 46, p659-66.

Sheares, B.T. and Robbins, P.W. (1986) Glycosylation of ovalbumin in a heterologous cell: analysis of oligosaccharide chains of the cloned glycoprotein in mouse L cells. Proc. Natl. Acad. Sci. USA, 83, p1993-1997.

Shimotohno, K, Kodama, Y., Hashimoto, J. and Miura, K.I. (1977) Importance of 5'-terminal blocking structure to stabilise mRNA in eucaryotic protein synthesis. Proc. Natl. Acad. Sci. USA, 74, p2734-2738.

Sigler, P.B. (1988) Acid blobs and negative noodles. Nature, 333, p210-212.

Simonsen, C.C. and Levinson, A.D. (1983) Isolation and expression of altered mouse dihydrofolate reductase cDNA. Proc. Natl. Acad. Sci. USA, 80, p2495-2499.

Sitia, R., Neuberger, M.S. and Milstein, C. (1987) Regulation of membrane IgM expression in secretory B cells: translational and post-translational events. EMBO J., 6, p3969-3977.

Smith, G.E., Frase, M.J. and Summers, M.D. (1983a) Molecular engineering of the *Autographa californica* nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene. J. Virol., 46, p584-593.

Smith, G.E., Summers, M.D. and Fraser, M.J. (1983b) Production of human beta interferon in insect cells infected with a baculovirus expression vector. Mol. Cell Biol., 3, p2156-2165.

Smith, G.L. and Moss, B. (1983) Infectious poxvirus vectors have capacity for at least 25000 base pairs of foreign DNA. Gene, 25, p21-28.

Smith, G.L., Mackett, M. and Moss, B. (1983c) Infectious vaccinia virus recombinant that expresses hepatitis B virus surface antigen. Nature, 302, p490-495.

Smithies, O., Gregg, R.G., Boggs, S.J., Koraleivski, A.M. and Kucherlapti, R.S. (1985) Insertion of DNA sequences into a human chromosomal β globin locus by homologous recombination. Nature, 317, p230-234.

Sompayrac, L.M. and Danna, K.J. (1981) Efficient infection of monkey cells with DNA of simian virus 40. Proc. Natl. Acad. Sci. USA, 78, p7575-7578.

Sorge, J., Cutting, A.E., Erdman, V.D. and Gautsch, J.W. (1984) Integration-specific retrovirus expression in embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA, 81, p6627-6631.

Southern, E.M., Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol., 98, p503-518.

Southern, P.J. and Berg, P. (1982) Transformation of mammalian cells to antibiotic reistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet., 1, p327-341.

Spier, R. (1988) Animal cells in culture: moving into the exponential phase. Trends in Biotech., 6, p2-6.

Sproat, B.S. and Bannwart, W. (1983) Improved synthesis of oligodeoxyribonucleotides on controlled pore glass using phosphotriester chemistry and a flow system. Tetrahedron Lett., 24, p5771-5774.

St. John, T.P. and Davis, R.W. (1981) The organisation and transcription of the galactose gene-cluster of *Saccharomyces*. J. Mol. Biol., 152, p285-315.

Stenflo, J., Fenlund, P., Egan, W. and Roepstorff, P. (1974) Vitamin K dependent modifications of glutamic acid residues in prothrombin. Proc. Natl. Acad. Sci. USA, 71, p2730-2733.

Sugden, B., Marsh, K. and Yates, J. (1985) A vector that replicates as a plasmid and can be effectively selected in B-lymphoblasts transformed by Epstein-Barr virus. Mol. Cell Biol., 5, p410-413.

Svasti, J. and Milstein, C. (1972) The complete amino acid sequence of a mouse κ light chain. Biochem. J., 128, p427-444.

Szybalska, E.H. and Szybalski, W. (1962) Genetics of human cell, IV. DNA-mediated heritable transformation of a biochemical trait. Proc. Natl. Acad. Sci. USA, 48, p2026-2034.

Talmadgo, K. and Gilbert, W. (1982) Cellular location affects protein stability in *E. coli*. Proc. Natl. Acad. Sci. USA, 79, p1830-1833.

Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. and Chambon, P. (1986) Requirement of stereospecific alignments for initiation from the SV40 early promoter. Nature, 319, p121-126.

Taketo, M., Gilboa, E. and Sherman, M.I. (1985) Isolation of embryonal carcinoma cell lines that express integrated recombinant genes flanked by Moloney murine leukemia virus LTR. Proc. Natl. Acad. Sci. USA, 82, p2422-2426.

Tartakoff, A.M. and Vassalli, P. (1978) Comparative studies of the intracellular transport of secretory proteins. J. Cell Biol., 79, p694-707.

Tataglia, J. and Paoletti, E. (1988) Recombinant vaccinia virus vaccines. Trends in Biotech., 6, p43-46.

Thummel, C., Tjian, R., Hu, S.-L. and Grodzicker, T. (1983) Translational control of SV40 T antigen expressed from the Adenovirus late promoter. Cell, 33, p455-464.

Tonegawa, S. (1983) Somatic generation of antibody diversity. Nature, 302 p575-581.

Treisman, R. (1985) Transient accumulation of c-*fos* RNA following serum accumulation requires a conserved 5' element and c-*fos* 3' sequences. Cell, 42, p889-902.

Tuite, M.F., Dobson, M.J., Roberts, N.A., King, R.M., Burke, D.C., Kingsman, S.M. and Kingsman, A.J. (1982) Regulated high efficiency expression of human interferon-alpha in *Saccharomyces cerevisiae*. EMBO J., 1, p603-608.

Tyler, B.M., Cowman, A.F., Gerondakis, S.D., Adams, J.M. and Bernard, O. (1982) mRNA for surface Ig γ chains encodes a highly conserved transmembrane sequence and a 28-residue intracellular domain. Proc. Natl. Acad. Sci. USA, 79, p2008-2012.

Uhlén, M., Moks, T. and Abrahmsén, L. (1988) Secretion of recombinant proteins into the culture medium of *E. coli* and *S. aureus*. Presented at the Biochem. Soc. meeting, Nottingham, England.

Ullrich, A. and Ramachandran, J. (1987) Insulin and type I insulin-like growth factor receptors. In 'The Molecular Biology of Receptors' (Ed. Strosberg, A.D.), p35-48, VCH Verlagsgellschaft, W. Germany.

Van Brunt, J. and Klausner, A. (1988) Growth factors speed wound healing. Biotechnology, 6, p25-30.

Van Eekelen, C.A.G. and van Venrooij, W.J. (1981) hnRNA and its attachment to a nuclear protein matrix. J. Cell Biol., 88, p554-563.

Vieira, J. and Messing, J. (1982) The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with universal primers. Gene, 19, p259-268.

Victor Garcia, J., Bich-Thuy, L., Stafford, J. and Queen, C. (1986) Synergism between Ig enhancers and promoters. Nature, 322, p383-385.

Vogelstein B. and Gillespie D. (1979) Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA, 76, p615-619.

Volloch, V. and Housman, D. (1981) Stability of globin mRNA in terminally differentiating murine erythroleukemia cells. Cell, 23, p509-514.

Voss, S.D, Scholkat, U. and Gruss, P. (1986) The role of enhancers in cell-type specific transcriptional control. Trends in Biochem. Sci., 11, p287-289.

Wabl, M.R. and Burrows, M.D. (1984) Expression of IgH chain at a high level in the absence of a proposed Ig enhancer element in *cis*. Proc. Natl. Acad. Sci. USA, 81, p2452-2455.

Wahl, G.M., de Saint Vincent, B.R. and DeRose, M.L. (1984) Effect of chromosomal position on amplification of transfected genes in animal cells. Nature, 307, p516-520.

Walter, P. and Blobel, G.J. (1981) Translocation of proteins across the endoplasmic reticulum. III. Signal recognition protein (SRP) causes signal-sequence - dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J. Cell Biol., 91, p557-561.

Walter, P. (1987) Two receptors act sequentially. Nature, 328, p763-764.

Wang, X.-F. and Calame, K. (1985) The endogenous IgH chain enhancer can activate tandem V_H promoters separated by large distance. Cell, 43, p659-665.

Ward, D. and Murray, N. (1979) Convergent transcription in bacteriophage λ: interference with gene expression. J. Mol. Biol., 133, p249-266.

Wasylyk, B., Wasylyk, C., Augereau, P. and Chambon, P. (1983) The SV40 72bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. Cell, 32, p503-514.

Wasylyk, B., Wasylyk, C. and Chambon, P. (1984) Short and long range activation by the SV40 enhancer. Nucl. Acids Res., 12, p5589-5608.

Wasylyk, B. and Wasylyk, C. (1986) The IgH-chain B-lymphoctyte enhancer efficiently stimulates transcription in non-lymphoid cells. EMBO J., 5, p553-56.

Weidle, U.H. and Buckel, P. (1987) Establishment of stable myeloma cells constitutively secreting human tissue-type plasminogen activator. Gene, 57, p131-141.

Weiher, H., Konig, M. and Gruss, P. (1983) Multiple point mutations affecting the simian virus 40 enhancer. Science, 219, p626-638.

Weinberger, J., Baltimore, D. and Sharp, P.A. (1986) Distinct factors bind to apparently homologous sequences in the IgH chain enhancer. Nature, 322, p846-848.

Weinberger, J., Jat, P.S. and Sharp, P.A. (1988) Localisation of a repressive sequence contributing to B-cell specificity in the IgH enhancer. Mol. Cell Biol., 8, p988-992.

Weintraub, H. (1985) Assembly and propagation of repressed and derepressed chromosomal states. Cell, 42, p705-711.

Weintraub, H., Feng Cheng, P. and Conrad, K. (1986) Expression of transfected DNA depends on DNA topology. Cell, 46, p115-122.

Whitelaw, E. and Proudfoot, N.J. (1986) α-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human α_2 globin gene. EMBO J., 5, p2915-2922.

Wiedmann, M., Kurzchalia, T.V., Hartmann, E. and Rapoport, T.A. (1987) A signal receptor in the endoplasmic reticulum membrane. Nature, 328, p830-833.

Wieland, F.T., Gleason, M.L., Serafini, T.A. and Rothman, J.E. (1987) The rate of bulk flow from the endoplasmic reticulum to the cell surface. Cell, 50, p289-300.

Wigler, M., Silverstein, S., Lee, L., Pellicer, A., Cheng, T. and Axel, R. (1977) Transfer of purified herpes thymidine kinase gene to cultured mouse cells. Cell, 11, p223-232.

Wigler, M., Sweet, R., Sim, G.K., Wold, B., Pellicer, A., Cheng, Y.C. and Axel, R. (1979) Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell, 16, p777-785.

Wigler, M., Perucho, M., Kurtz, D., Dana, S., Pellicer, A., Axel, R. and Silverstein, S. (1980) Transformation of mammalian cells with amplifiable dominant-acting gene. Proc. Natl. Acad. Sci. USA, 77, p3567-3570.

Williams, D.C., Frank van, R.M., Muth, W.L. and Burnett, J.P. (1982) Cytoplasmic inclusion bodies in *E. coli* producing biosynthetic human insulin proteins. Science, 215, p687-688.

Williams, G. (1988) Novel antibody reagents: production and potential. Trends in Biotech., 6, p36-42.

Williams, J.G. and Mason, P.J. (1985) Hybridisation in RNA analysis. In 'Nucleic acid hybridisation - a practical approach' (eds. Hames B.D. and Higgins S.J.), chapter 6; p139-160, IRL Press; Oxford, England and Washington D.C., USA.

Wingender, E. (1983) Compilation of transcription regulating proteins. Nucl. Acids Res., 16, p1879-1902.

Wold, F. (1981) *In vivo* chemical modification of proteins. Ann. Rev. Biochem., 50, p738-814.

Wu, R.S., Pansuz, H.T., Hatch, C.L. and Bonner, W.M. (1986) Histones and their modifications. CRC Crit. Rev. Biochem., 20, p201-263.

Wurm, F.M., Gwinn, K.A. and Kingston, R.E. (1986) Inducible overproduction of the mouse c-*myc* protein in mammalian cells. Proc. Natl. Acad. Sci. USA, 83, p5414-5418.

Yang, L., Rowe, T.C., Nelson, E.M. and Liu, L.-F. (1985) *In vivo* mapping of topoisomerase II specific cleavage sites in SV40 chromatin. Cell, 41, p127-132.

Yancopoulus, G.D. and Alt, F.W. Developmentally controlled expression of unrearranged V_H gene segments. Cell, 40, p271-281.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene, 33, p103-109.

Yaniv, M. and Cereghini, S. (1986) Structure of transcriptionally active chromatin. CRC Crit. Rev. Biochem., 21, p1-26.

Yates, J., Warren, N., Reisman, D. and Sugden, B. (1984) A *cis*-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc. Natl. Acad. Sci. USA, 81, p3806-3810.

Yen, T.J., Machlin, P.S. and Cleveland, D.W. (1988) Autoregulated instability of β-tubulin mRNAs by recognition of the nascent amino terminus of β-tubulin. Nature, 334, p580-585.

Yoshimura, F.K., Davison, B. and Chaffin, K. (1987) Different activities of viral enhancer elements before and after stable integration of transfected DNAs. Mol. Cell Biol., 7, p1296-1299.

Young, R.A. and Davis, R. (1983) Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA, 80, p1194-1198.

Yuan, D. and Tucker, P.W. (1984) Transcriptional regulation of the mu-delta heavy chain locus in normal murine lymphocytes-B. J. Exp. Med., 160, p564-583.

Zaller, D.M. and Eckhardt, C.A. (1985) Deletion of a B cell specific enhancer affects transfected, but not endogenous IgH chain enhancer expression. Proc. Natl. Acad. Sci. USA, 82, p5088-5092.

Zerler, B., Moran, B., Maruyama, K., Moomaw, J., Grodzicker, T. and Ruley, H.E. (1986) Adenovirus E1A coding sequences that enable *ras* and *pmt* oncogenes to transform cultured primary cells. Mol. Cell Biol., 6, p887-899.

ABSTRACT

Development of Vectors Allowing Efficient Heterologous-Gene expression in Stable Myeloma-Cell Transfectants.

Kevin Hudson; Ph.D. Thesis. Department of Biochemistry, University of Leicester, University Rd., Leicester. LE1 7RH. England.

This thesis describes the development of expression-plasmid vectors for use in the mouse, myeloma cell-line, J558L. Following introduction of the plasmids into J558L using spheroplast fusion, the vectors allow selection for stably-transfected cells. Selective survival depends on expression from a *gpt* gene in the plasmid, which encodes the enzyme, xanthine-guanine phosphoribosyltransferase (XGPRT). Expression of XGPRT confers, on J558L, a dominant selectable-resistance to mycophenolic acid in the presence of xanthine and hypoxanthine. The resultant stable *gpt*$^+$-transfectants can then be screened for expression of a protein encoded by a non-selected gene in the expression plasmid.

Initially, a systematic study was carried out to compare the effect on XGPRT expression of different expression elements placed upstream of the *gpt* coding sequence, in single-gene plasmids. Comparative expression levels were estimated by comparing the stable *gpt*$^+$-transfection frequencies of J558L obtained with each plasmid. This study demonstrated the importance of the IgH-gene enhancer for obtaining high-level expression from a transfected gene. A combination of an IgH enhancer and a promoter element, upstream of the *gpt* coding-sequence, resulted in the highest levels of expression, but the type of promoter used was of only secondary importance.

Identified combinations of effective upstream elements were then incorporated into plasmid vectors, which were constructed for expression from heterologous genes encoding proteins of interest. Stable *gpt*$^+$-transfectants containing a non-selected gene were only obtained following transfection with plasmids which also contained the *gpt* gene. This was in contrast to situations in which a non-selected gene was cotransfected with a *gpt* gene which resided on a different plasmid; here, none of the stable *gpt*$^+$-transfectants screened had cointegrated the non-selected gene with the *gpt* gene. Different classes of the two-gene expression plasmids were constructed, which differed in the relative orientation and position of the transcription units. The efficiency of the various plasmids was estimated by measuring the expression levels from a model cDNA, encoding chicken lysozyme. Transfectants were isolated which express and secrete biologically-active chicken lysozyme, at levels greater, in molar terms, than the typical level of secretion of endogenous Ig from myeloma cells.

In some cases, the stable *gpt*$^+$-transfection frequencies obtained with two-gene plasmids were low, compared with the transfection frequencies obtained with plasmids containing only the *gpt* gene. It was proposed that this was due to transcriptional interference between the two genes on the same plasmid. As this phenomenon might also restrict the expression from the non-selected gene, overcoming this effect was considered important in optimising expression. However, attempts to identify and overcome the phenomenon were inconclusive and, consequently, a rational strategy for overcoming this potential limitation on expression levels was not obtained.