Expression of Microbial Cysteine Biosynthesis Genes in Sheep by Transgenesis.

Christopher Simon Bawden
B.Sc.(Hons.), Grad. Dip. Ed..

A thesis submitted for the degree of
Doctor of Philosophy in the University of Adelaide
Department of Biochemistry
March, 1991.

For my dear Carolyn,

for my children, Candice, Joel and Timothy,

and for my parents.

Thesis Summary

The main body of work undertaken in this thesis involves analysis of expression of the cysteine biosynthesis genes of the yeast, Saccharomyces cerevisiae, and the bacterium, Salmonella typhimurium, in various mammalian systems including cell cultures, transgenic mice and transgenic sheep.

The ultimate goal of the work is to establish a heritable trait of increased wool growth in the sheep by transgenesis with the genes for cysteine biosynthesis. This has arisen from a knowledge of the limitation imposed on wool growth by an insufficient supply of the sulphur-containing amino acid cysteine or its precursor, methionine, to the wool follicle. This limitation has been well established experimentally, and is a result of many factors, not the least of which is a requirement for large quantities of cysteine to support synthesis of the keratin proteins of the wool fibre.

Since the physiological mechanisms involved with cysteine metabolism in the rumen of the sheep limit wool growth, a logical solution would be to provide the sheep with a capacity to synthesize larger quantities of cysteine endogenously. This could be achieved by transgenesis of the sheep with a set of cysteine biosynthesis genes not found in the mammalian genome. In contrast to mammals, plants and microbial species can synthesize cysteine in a two-step reaction. The microbial enzymes of this pathway, serine acetyltransferase and O-acetyl L-serine sulphydrylase, have been well characterised. The genes encoding the respective enzymes in yeast and bacteria are CYS1 and MET25 (yeast) and cysE and cysM (bacteria). The latter three of these have been isolated and positively identified.

Testing the fidelity of expression of the cysteine synthesis genes from the species S. cerevisiae and S. typhimurium in a mammalian environment is a prerequisite to their use in such a transgenesis programme. This thesis first describes analysis of the expression of the yeast gene, MET25, in transient and permanent mammalian cell culture systems, and finds it unsuitable as a transgene. Subsequently, the analysis of expression of the cysE gene from Salmonella in the same cell culture systems and transgenic mice is described. Detection of tissue-specific expression of functional enzyme molecules in vivo (transgenic mice) demonstrated this to be a suitable candidate for transgenesis of the sheep.

Finally, analysis of expression of a combination of the cysE and cysM genes from Salmonella in a permanent cell culture system, in transgenic mice and in transgenic sheep, is described. Determination of expression of both genes in cloned cell lines, through detection of the desired catalytic conversion of serine to cysteine, revealed a capacity to synthesise cysteine. Following this, the expression of the same gene combination in transgenic mice and sheep was analysed.

The provision of a mammalian species with a new amino acid synthesis pathway would indeed be a unique situation, purely from the point of view of the presence of such an extra metabolic option within a cell. The consequences of the availability of such a pathway to a normal mammalian cell is unknown but will become apparent at the stage of testing the cysteine biosynthesis genes in a tissue-specific manner in the epithelial cells lining the rumen, where the vital substrate sulphide is found in abundance. In the final chapter, this thesis discusses the possible ramifications of such an achievement for the physiology of the sheep.

Declaration

This thesis contains no material which has been accepted for the award of a degree or diploma in any University and in my belief, contains no material which has been published by another person, except where due reference is given.

Christopher Simon Bawden

NAME: ...C. S. Bawden... COURSE: ...Biochemistry, Ph.D...

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED: .... DATE: ...15/10/91...

Acknowledgements

Abbreviations

acetyl CoA :	acetyl Coenzyme A
bp :	base pair
BSA :	bovine serum albumin
CHO :	Chinese Hamster Ovary
Ci :	curie
cys-4VP :	4-pyridyl-ethyl cysteine
dATP :	deoxyadenosine triphosphate
dCTP :	deoxycytidine triphosphate
dNTP :	deoxynucleoside triphosphate
DNA :	deoxyribonucleic acid
E. coli :	Escherichia coli
F_0 :	founder generation
F_1 :	first filial generation
F_2 :	second filial generation
hGH :	human Growth Hormone
kb :	kilobase pair
kD :	kiloDalton
M_r :	molecular weight
mRNA :	messenger RNA
NAS :	N-acetyl-L-serine
OAS :	O-acetyl-L-serine
PAGE :	polyacrylamide gel electrophoresis
RNA :	ribonucleic acid
RSVLTR :	Rous sarcoma virus long terminal repeat
rUTP :	ribouridine triphosphate
SDS :	sodium dodecyl sulphate
S cerevisiae :	Saccharomyces cerevisiae
S.typhimurium :	Salmonella typhimurium
SV40 :	Simian virus 40
TLC :	thin layer chromatography
Tris :	Tris (hydroxymethyl) aminomethane
tRNA :	transfer RNA

Contents

Chapter 1

Introduction.

Chapter 1 : Introduction.

1.1 General Introduction.

A longstanding desire to improve the phenotype of the existing Australian sheep breeds has been the driving force responsible for many of the genetic improvement programmes mounted to date. The aim of these programmes has been to overcome the limitations imposed upon wool growth by various genetic and environmental factors, such that a finer quality and greater quantity of fleece is obtained. These have succeeded to the point at which the current Australian fine wool Merino sheep produces a fleece from about one hundred million follicles. A critical balance between many nutritional parameters must be maintained to sustain such wool growth, and even if this is achieved, wool growth is known to be sub-optimal due to an inadequate supply of the sulphur-containing amino acid, cysteine, to the actively synthesising wool follicle. This is recognised as the primary limitation placed upon wool growth. Whilst no amount of effort in selective breeding could overcome this problem, it lends itself to the technologies of genetic engineering and transgenesis, through which it is possible to direct phenotypic alterations via transfer of genetic information from one species to another. The nutritional inadequacy which limits wool growth could perhaps be solved by providing the sheep with a capacity to synthesize greater quantities of its own cysteine, through transgenesis of specific genes involved with cysteine synthesis.

Genes encoding enzymes responsible for a two-step de novo synthesis of cysteine exist in the genomes of plants, fungi and microorganisms and the protein products of the genes from yeast and bacteria have been extensively studied. If the appropriate genes could be isolated, transferred to the sheep and expressed at an adequate level and suitable site, such that wool follicles could be supplied directly with enough cysteine to meet their synthesis needs, then an increase in wool growth in those transgenic sheep would be expected. This experimental approach, utilising the cysteine synthesis genes from the yeast, Saccharomyces cerevisiae and the bacterium, Salmonella typhimurium forms the basis of the project to be described in this thesis.

The remainder of the present chapter gives a more detailed account of the requirement for cysteine in wool growth including structural considerations of the wool fibre itself, the evidence which determines the relationship between cysteine and sub-optimal wool growth and the factors which contribute to the sulphur deficit in sheep. Other genetic and environmental factors known to limit wool growth will only be considered here in the context of their relationship to the limitation placed on wool growth by an inadequate supply of cysteine. The cysteine biosynthesis pathways of the microorganisms S. cerevisiae and S. typhimurium will also be discussed.

1.2 The requirement for cysteine in wool growth.

As shown by abundant evidence gathered from experiments with Australian sheep breeds, both in natural and penned grazing environments, normal wool growth is sub-optimal. Many factors, some genetic and some environmental, are known to be responsible for this (for reviews, see Black and Reis, 1979; McGuirk, 1987). Amongst other nutritional concerns, it had been suspected for some time that a major factor limiting the growth rate of wool was an inadequacy in the supply of the sulphur-containing amino acid, cysteine, or its precursor methionine, to the wool follicle. This notion, supported in part by the fact that plant proteins in the sheep diet are generally poor in the sulphur-amino acids, seems quite logical if one considers the amino acid composition of the keratin proteins of the wool fibres themselves as dictated by the structural requirements for rigidity and strength (via disulphide crosslinking of proteins) of the fibre.

A wool fibre is comprised of two main structural keratin protein groups known as the intermediate filament proteins (IF) and intermediate filament associated proteins (IFAP) (Powell and Rogers, 1986; Rogers et al., 1989) found in approximately equal proportions in the wool fibre. The IF proteins which associate into axially-oriented microfibrils in the cortical cells of the developing fibre, contain low proportions of cysteine (approximately 6 moles %) and are known as the low-sulphur keratin proteins. The proteins vary in size from about 350 - 550 amino acids and contain 20 - 30 cysteine residues. In the case of the

high-sulphur and ultra-high sulphur keratin proteins of the IFAP group, which hold adjacent filaments rigidly in place, cysteine comprises between 16 - 24 moles %, and in excess of 30 moles % , of the proteins respectively (Gillespie and Broad, 1972; Crewther, 1976). For these proteins which vary from 99 - 180 amino acids in size, this equates to 20 - 60 cysteine residues per molecule. As a result, the mole percent of cysteine in total wool protein is quite high at around 13 moles % (Marshall and Gillespie, 1977).

A quantitative assessment of the requirement for cysteine in wool growth is found in the results of a study made by Hogan et al. (1979) of maximum wool growth rates in various Australian Merino genotypes. Under carefully controlled experimental conditions, in which the feed intake was monitored with respect to the amount of dietary cysteine available and actually absorbed by the sheep, wool growth rates of 10.3 grams/day (Saxon) to 22.0 grams/day (Strong South Australian) were observed. These are to be compared with industry estimates of 5.9 grams/day (Saxon) and 12.7 grams/day (Strong South Australian) for ewes in commercial flocks grazed under favourable conditions. The daily output of cysteine in the wool produced by these different genotypes was also measured. In the case of the Saxon merino, this was found to be 0.97 - 1.49 grams/day. The dietary cysteine initially absorbed was 2.28 grams/day, so this represents only 42.5 - 65.4 % utilisation of the total available cysteine. The output of cysteine in wool of the Strong South Australian merino was 2.27 - 3.48 grams/day and since the dietary cysteine initially absorbed was 4.04 grams/day, 56.2 - 86.1 % of the total available cysteine was utilised. Clearly, cysteine, as a major component in the protein structure of wool is required in large quantities to sustain wool growth.

The other sulphur-containing amino acid, methionine, is a precursor for cysteine synthesis. As an essential amino acid, it must be supplied in the diet to support synthesis of all proteins, including keratin protein synthesis in the wool fibre via the production of cysteine in the transulphuration pathway (to be discussed further in 1.4). Other effects upon wool growth which have been attributed to these amino acids include the requirement for methionine in polyamine synthesis (Reis, 1989) and the stimulation of mitosis in the follicle bulb by increased levels of cysteine or methionine (Hynd, 1989). Obviously,

growth and formation of wool in the estimated 100 million follicles of a sheep would be adversely affected by a continuous shortfall in these amino acids, and wool growth would indeed be sub-optimal.

1.3 The relationship between cysteine and sub-optimal wool growth.

The first experimental evidence which established that an inadequate supply of the sulphur-containing amino acids severely limits the rate of wool growth was provided by the work of Reis and Schinkel (1963, 1964). It was clearly demonstrated that if cysteine or methionine were administered by abomasal infusion at a rate of 1.5 - 3.0 grams/day, wool growth was substantially increased. (Even as little as 0.5 grams/day of cysteine was observed to elicit a response of increased wool growth; Reis, 1967). Whilst the response to such treatment is seen to vary between genotypes (Hogan et al., 1979; Williams, 1979) and also to be subject to the general nutritional plane of the individual sheep (Allden, 1979), an astounding 100 % increase in wool growth rate was observed in some cases. This was manifested as an increase in fibre volume (40 - 50%) with increases in both length growth rate and fibre diameter (Reis et al., 1973a; Reis, 1989).

Many changes have been documented at the cellular and molecular level which illustrate the capacity of the wool follicle to utilize larger quantities of cysteine. At the molecular level, specific alterations in the protein composition of the wool fibre are observed; namely an increase in the level of the ultra-high sulphur IFAP keratins (Gillespie et al., 1964; Gillespie et al., 1969) and a decrease in the level of the high-glycine/tyrosine IFAP keratins (Frenkel et al., 1974; Frenkel et al., 1975). There are associated changes in the ratio of these IFAPs to one another and in their absolute amounts in the wool fibre. Although the overall sulphur content of the fibre increases, no increase is observed in the other high-sulphur IFAPs. Similarly, there is no increase in the IF keratin proteins (Gillespie et al., 1969). The molecular mechanisms of these changes and their significance for wool fibre properties is unknown. However to improve the present understanding of the response of the wool follicle to increased levels of cysteine, changes in the level and stability of specific follicular IF and IFAP

mRNA molecules upon cysteine infusion is currently being investigated. It is interesting to note that infusion of no other single amino acid, per abomasum, produced an effect such as that seen upon administration of cysteine or methionine (Reis, 1979).

The findings of Reis and Schinckel (1963, 1964) seemed in direct conflict with the earlier observation that dietary supplements of cysteine or methionine given orally (c.f. abomasal infusion) failed to stimulate wool growth. This, however, was reconciled when more sophisticated knowledge of the ruminant stomach, its physiology, and its relationship to the site of infusion became available. The sheep is a ruminant animal with a stomach consisting of four main compartments. The first three of these, the rumen, reticulum and omasum, in order of their proximity to the oesophagus, comprise the forestomach, and are lined with a nonglandular epithelium. The rumen is the largest of the three and contains an abundant microflora. The fourth compartment is the abomasum and is lined with a glandular epithelium. Food subjected to mechanical breakdown in the mouth is degraded further upon entering the rumen by the chemical action of the microbes present there (including bacteria, fungi and protozoa). These microbes particularly attack the cellulose of the cell walls of the plant material ingested, and so derive their own energy from this, and from any other substances including proteins released from damaged plant cells. A major product of the fermentative action of the microbes in the rumen is the volatile fatty acids, which are absorbed from this compartment and used as an important energy source by the sheep.

Once in the abomasum, the digestive chemistry of the sheep comes into play and the liquid digesta, including dead microbes, is more thoroughly degraded. It is from this compartment, and the small intestine, which the sheep derives nutrient amino acids by absorption directly into the bloodstream (figure 1.1). Prior to reaching the abomasum, however, extra cysteine or methionine given orally in the diet is first lost to the microbes of the rumen. On the surface, this may appear to explain why such dietary supplementation with cysteine or methionine has no effect on the rate of wool growth. A more complete understanding of the reason for a shortfall in sulphur-containing amino acids in the follicle, indeed a sulphur deficit at this site, requires a closer examination of rumen physiology.

Figure 1.1 : Digestion and metabolism of plant protein and carbohydrate in the sheep rumen and abomasum in relation to wool growth.

Mechanical breakdown of food in the mouth is followed by degradation in the rumen and abomasum. Carbohydrate is converted to volatile fatty acids (VFA) in the rumen (1). The VFA's enter the blood (2) and form a major energy source for the sheep. Dietary protein is first broken down into its constituent amino acids (3). Some of these are assimilated into microbial protein (4) which is then degraded in the abomasum. Amino acids released from the microbial protein pass from the abomasum to the small intestine and then to the bloodstream (5). From here, they are taken for protein synthesis in the sheep, including synthesis of keratin proteins in the wool follicle. Dietary amino acids not utilised for microbial protein synthesis are degraded to elemental compounds. In the case of the sulphur amino acids, cysteine and methionine, H_2S is produced (6) and passes rapidly from the rumen via the epithelial lining to the bloodstream (7). Much of the sulphide which enters the blood is converted to sulphate and excreted in the urine. The diagram is adapted from Rogers, 1990.

Plant Protein
Carbohydrate

10^8 Wool Follicles

2.

Blood
Pool

Cellulose Hydrolysis
and Fermentation
to VFA

1.

RUMEN

7.

5.

6.

ABOMASUM

H$_2$S

Amino Acids

INTESTINE

Amino Acids → Microbial
Protein

3.

4.

Plant Proteins
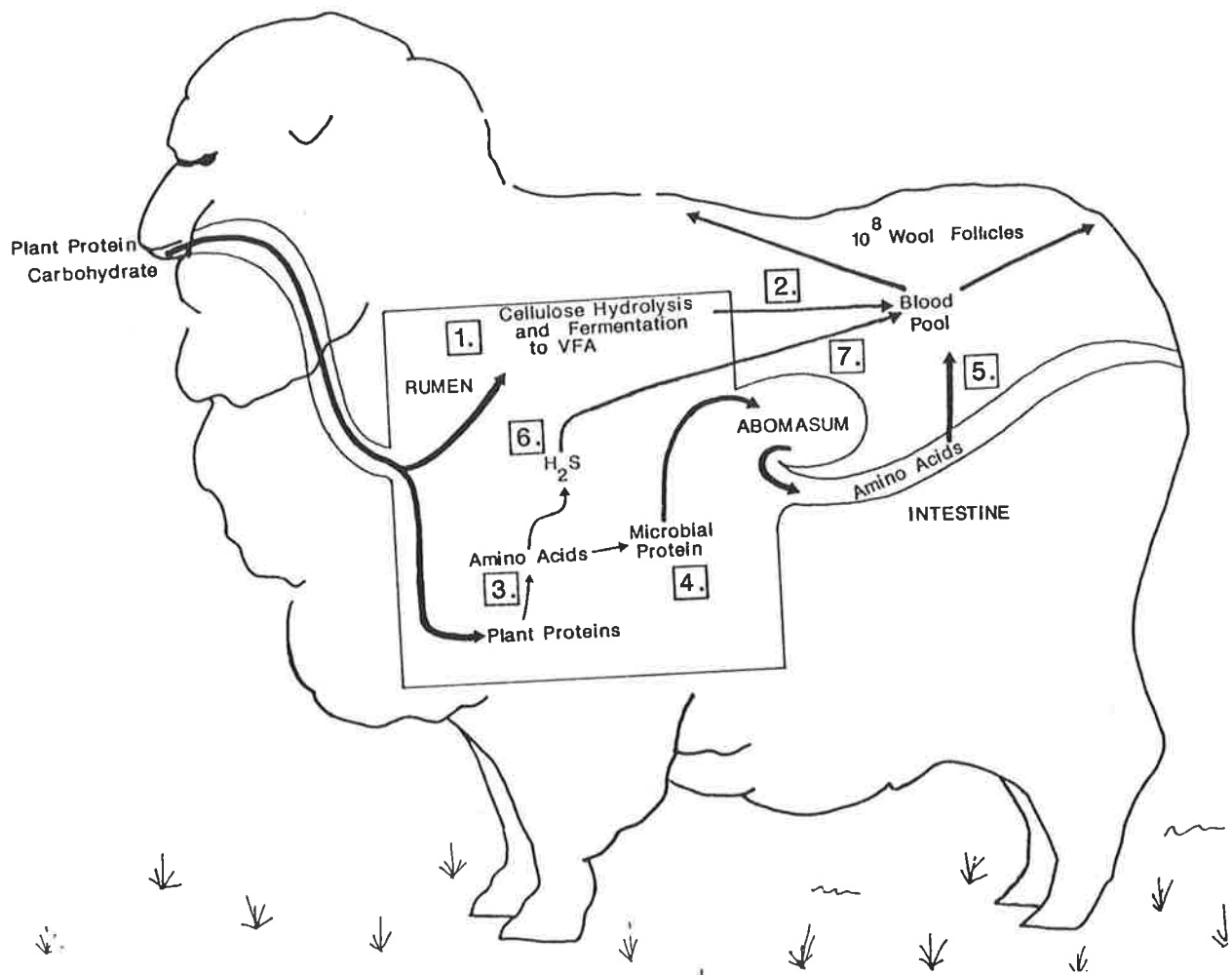
1.4 Factors which contribute to a sulphur deficit in the wool follicle.

One might expect that dietary supplements of cysteine and methionine given orally, if incorporated into extra bacterial protein, would subsequently be reclaimed by the degradative processes of the abomasum, acting on the microbes themselves. This is the usual route of dietary cysteine and methionine into the bloodstream (Faichney and Black, 1979; Cheng et al., 1979). However, like those released from plant protein by the proteases of rumen microbes, the sulphur-containing amino acids given as dietary supplements may suffer a different fate, namely degradation, to release the sulphur as sulphide and finally sulphite and sulphate (Lewis, 1954; Henderickx, 1961). Experimentation with ^{35}S cysteine and methionine revealed that these amino acids are extensively degraded in the rumen, cysteine more rapidly than methionine (Bird, 1972). Following such extensive degradation, there is a rapid flux of the major end product, sulphide, into the bloodstream via the epithelial cells which line the rumen. To illustrate this, Bray measured the half-life of ^{35}S sulphide infused directly into the rumen and found it to be very short, an estimated 10-20 minutes (Bray, 1969a). This is an important fact and will be discussed further in section 1.7. Oxidation of the sulphide to sulphate in the blood and the liver, is followed by excretion of some of the sulphate from the body of the sheep in the urine. Measured directly by Bray in the same experiment as detailed above, 30 % of the initial dose of sulphur (^{35}S) given to the rumen as sulphide is found in the urine 3 hours post-infusion (Bray, 1969a; also see Bray, 1969b and Bray and Till, 1975).

In a separate study by Kennedy and Milligan (1978) the amount of inorganic sulphur excreted in the urine of sheep receiving different amounts of sulphur in their diet was measured and seen to vary considerably; 21 - 68 milligrams/day for sheep receiving brome grass (1.44 grams of sulphur/kilogram of dry feed) and 427 - 956 milligrams/day for sheep receiving lucerne (2.14 grams of sulphur/kilogram of dry feed). In those receiving lucerne, the corresponding intake of sulphur in the feed was over the range of 1755 - 3510 milligrams/day. In the sheep at these extremes, 24.3 % (i.e. 427 mg from 1755

mg) and 27.2 % (i.e. 956 mg from 3510 mg) of the mass of sulphur obtained from the feed was excreted from the sheep as sulphate in the urine each day.

These results demonstrate the net loss of sulphur from the sheep, a factor which contributes greatly to the "sulphur deficit" of the follicles. In the latter case, it is interesting to note that the amount of sulphur lost in the urine each day is equivalent to that which might be found in 1.62 - 3.62 grams of cysteine, a range corresponding directly to the levels adopted for abomasal infusion by Reis and Schinckel (1963, 1964) of 1.5 - 3.0 grams of cysteine per day. In addition, according to the measurement of rates of wool growth and utilisation of cysteine in Merino sheep (Hogan et al., 1979) the net urinary loss of sulphur per day measured by Kennedy and Milligan, even considering the lower figure of 427 milligrams/day, would theoretically be equivalent to that in an extra 7.1 - 10.9 grams (Saxon) and 20.0 - 30.7 grams (Strong South Australian) of wool per day. Although these data are not strictly comparable, this calculation quantitatively highlights the contribution of microbial degradation of sulphur-amino acids to the sulphur deficit in the wool follicle and shows that the possible increase in wool growth rate is significant.

If the cysteine or methionine is given per abomasum, the loss of sulphur from the sheep by the degradative action of the bacteria and protozoa is avoided. The intact amino acids have a direct route to the bloodstream and subsequently to the wool follicle. Once again, the use of tracer doses of ^{35}S cysteine and methionine has clearly illustrated this, and has shown these amino acids to be incorporated into wool protein, as cysteine, with very high efficiency.(Downes et al., 1964; Downes et al., 1970). The definitive results of these experiments give direct evidence of the fate of infused cysteine and methionine.

In addition to those mentioned, another factor pertaining to the normal source of supply of cysteine to the wool follicle contributes to sub-optimal wool growth. Apart from the metabolic shuttling of endogenous, protein-bound cysteine, the other main source of supply of free cysteine to the wool follicle is the cysteine which is produced as an end product of the metabolism of methionine. The pathway involved is known as transulphuration, in which the last steps to cysteine synthesis are irreversible. In the sheep, this pathway of conversion of methionine to cysteine is known to be functional in several tissues,

particularly in the liver and kidney, and also in the skin (Radcliffe and Egan, 1974; Downes et al., 1964). All mammals rely on this pathway for the synthesis of endogenous cysteine. However, because methionine degradation in the rumen is extensive, the use of increased transulphuration as a means of improving the supply of cysteine to the wool follicle is not possible.

In summary, given the huge demand for sulphur-amino acids during the synthesis of proteins in the wool follicle, the main factors responsible for the sulphur deficit and resulting limitation to wool growth in the follicle are;

(a) low dietary intake of cysteine (and methionine) from feed protein,

(b) degradation of dietary cysteine and methionine by microbes in the rumen, and net loss of sulphur from the sheep (sulphide produced by the microbes is converted to sulphate and excreted from the sheep) and

(c) restriction of transulphuration function by the degradation of dietary methionine in the rumen.

1.5 The use of Genetic Engineering to overcome the sulphur deficit.

To overcome the factors which contribute to sub-optimal wool growth, and in particular the sulphur deficit described, either the environment or the genotype of the sheep must be altered. The obvious economic importance of improvements to wool production has given rise to a number of research efforts whose aim is to pursue one or other of these alternatives (for a review, see Rogers, 1990). In all cases, the methods of genetic engineering have become central to the success of the experiments being undertaken. For example, the microbes of the rumen are being genetically engineered to produce proteins that are resistant to degradation in the rumen and to contain a balance of amino acids beneficial to wool growth, such that hydrolysis of microbial protein in the abomasum releases higher levels of these amino acids for use by the sheep (Brooker et al., 1989). Another attempt towards alteration of the environment of the sheep, through plant transgenesis, is the enrichment of the sulphur content of the leaf protein of a feed plant such as lucerne by the introduction of foreign genes encoding proteins which are rich in cysteine and/or methionine. The proteins chosen, such as chick ovalbumin (W.R. Knibb, pers. comm.) and pea albumin 1

(Higgins et al.,1989) are resistant to degradation in the rumen. Post-ruminal digestion of these extraneous feed proteins would be expected to deliver more cysteine to the wool follicle via the normal dietary route.

Current and future alterations to the genotype of the sheep also depend on the use of genetic engineering. In contrast to this, the success of the genetic improvement programmes initiated to date has relied solely on conventional, selective breeding techniques. Phenotypic alterations such as those related to wool fibre diameter and crimp, to efficacy of feed usage and fecundity, have been achieved by selective breeding. These are only a few examples (McGuirk, 1987; Black and Reis, 1979). Though quite laborious, selective breeding has offered the most satisfactory solution to attainment of the desired phenotypes. A constraint, however, is that it may only draw upon the gene pool of the species(Ovis aries) in order to make a phenotypic improvement affording finer quality or greater quantity of fleece. Eventually, a point will be reached when enhancement of wool growth phenotype cannot be achieved by this conventional approach to genetic manipulation.

The limitation to wool growth described here presents a case which lends itself to application of the new technologies of genetic engineering and transgenesis. The problem of an inadequate supply of cysteine and methionine to the follicle cannot be overcome by manipulation of the genes of this species alone; the solution lies in the use of genes from a different species, in this case genes from a microbial source. It is clear that sub-optimal wool growth may be markedly improved by abomasal or intravenous infusion of the amino acid cysteine, at a rate of 1.5 - 3.0 grams/day. If the sheep itself was able to synthesize a quantity of cysteine equivalent to this by means other than transulphuration, then one might expect to observe a similar improvement in the growth rate of the wool.

In contrast to vertebrate animals, plants, fungi and bacteria all have the capacity to synthesize cysteine de novo in a simple two-step reaction. In the first reaction, L-serine is converted to O-acetyl L-serine (OAS) by the action of an acetyltransferase, which transfers the acetyl group of acetyl-CoA to serine (figure 1.2). A sulphydrylase then converts the OAS to L-cysteine. Sulphide is an essential substrate in this latter reaction step. The aim of this project, discussed in

Figure 1.2 : The microbial cysteine biosynthesis pathway.

L-serine is first converted to O-acetyl-L-serine by serine acetyltransferase (SAT). The intermediate, O-acetyl-L-serine, is then converted to L-cysteine by an O-acetyl-L-serine sulphydrylase enzyme (OAS sulphydrylase).

L-serine + acetyl CoA $\xrightarrow{\quad 1 \quad}$ O-acetyl-L-serine + CoA

O-acetyl-L-serine + H_2S $\xrightarrow{\quad 2 \quad}$ L-cysteine + acetate + H_2O
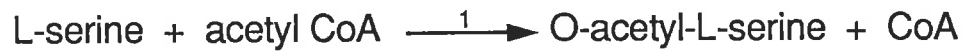
1. Serine acetyltransferase (SAT)

2. O-acetyl-L-serine sulphydrylase (OAS sulphydrylase)

detail in section 1.9, is to establish a heritable trait of increased wool growth by providing the sheep with the ability to synthesize optimal quantities of cysteine. In principle this could be achieved by transgenesis of the sheep with the two genes whose protein products would facilitate the enzymic conversion of L-serine to L-cysteine (see Bawden et al., 1987; D'Andrea et al., 1989, and Rogers, 1990).

1.6   The cysteine biosynthesis pathways of microorganisms.

In contrast to the higher eukaryote plant species, the cysteine biosynthesis pathways of the lower eukaryote organisms such as the yeast S. cerevisiae and the prokaryotes such as the bacterium S. typhimurium have been researched extensively. A great deal of data has been accumulated relating to many of the enzymes involved in these metabolic pathways, and to the genes which encode them. Many auxotrophic mutants have been discovered and used to define the catalytic steps in the respective pathways. For this reason, the cysteine biosynthesis genes of these particular species have been sought as candidate transgenes in the sheep transgenesis programme outlined. In each case, genes have been isolated by standard recombinant DNA methods. Considering information regarding the catalytic activity and physical properties of the enzymes in question, it is assumed that expression of various combinations of these yeast and bacterial genes would be possible in order to provide a complete cysteine biosynthesis pathway to the sheep. These genes will now be discussed in more detail, with reference to the cysteine biosynthesis pathways of these organisms, first the yeast, Saccharomyces cerevisiae then the bacterium, Salmonella typhimurium.

1.6.1   The cysteine biosynthesis genes of Saccharomyces cerevisiae.

In yeast and mammals, cysteine and methionine biosynthesis is directly related through the pathway known as transulphuration (de Robichon-Szulmajster and Cherest, 1967; Cherest et al., 1969; Cooper, 1983; Griffith, 1987; Soda, 1987). Differences exist between these pathways (figure 1.3), but in each case the precursor methionine may be converted to cysteine. One notable difference, the

Figure 1.3 : Pathways of cysteine and methionine biosynthesis, linked by transulphuration.

Figure 1.3 (a) : Cysteine and methionine biosynthesis in yeast. Genetic blocks for which the cys1/cys2 mutants (defective in the serine acetyltransferase conversion step; 1) and the met17/met25 mutants (defective in the O-acetyl-L-serine and O-acetyl-L-homoserine sulphydrylase steps; 2 and 3) are responsible, are indicated at the appropriate steps in the pathway. The two step synthesis of cysteine is shown in reactions 1 and 2. The entry point of sulphide (derived from the sulphate assimilation process) into the pathway is also shown. The pathway is adapted from that shown in D'Andrea et al., 1989.

Figure 1.3 (b) : Cysteine and methionine biosynthesis in mammals. The enzymes responsible for the interconversion of methionine and cysteine are shown below the reaction pathway. The final steps from serine to cysteine via cystathionine are irreversible. This pathway accounts for the cysteine which is synthesised from dietary methionine (e.g. as seen in the skin of the sheep). The pathway is from Reis, 1979.

a.

S-ADENOSYL-METHIONINE

METHIONINE

S-ADENOSYL-HOMOCYSTEINE

SERINE

TRANSULPHURATION PATHWAY

α-KETOBUTYRATE +NH₃

SO₄⁻

HOMOCYSTEINE

CYSTATHIONINE

CYSTEINE

SULPHATE ASSIMILATION

③

S⁻

PYRUVATE +NH₃

②

S⁻

O-ACETYLHOMOSERINE

O-ACETYL HOMOSERINE

O-ACETYLSERINE

HOMOSERINE

Co A

①

Ac Co A

THREONINE

SERINE

OXALOACETATE

b.

D-Methionine → L-Methionine → Protein synthesis / Chain initiation / Incorporation into protein

ATP

Enzyme 1

ADP+P$_i$

Methyl group donors

S-Adenosyl-L-methionine

Acceptors

Methylated acceptors

S-Adenosyl-L-homocysteine

Adenosine

L-Homocysteine

Enzyme 2

L-Serine

L-Cystathionine

Enzyme 3

α-Ketobutyrate

L-Cysteine ⇌ L-Cystine

Sulphate

Enzymes

1. Methionine adenosyltransferase (EC 2.5.1.6)
2. Cystathionine β-synthase (EC 4.2.1.22)
3. Cystathionine γ-lyase (EC 4.4.1.1)

inability of mammals to synthesize appreciable amounts of cysteine utilising a different precursor, serine (Cooper, 1983), is the basis of the current transgenesis programme. As stated previously, the synthesis of cysteine in microorganisms occurs in two stages (see figure 1.2). The first reaction involves the conversion of L-serine to O-acetyl L-serine (OAS) and in yeast the serine acetyltransferase (SAT) enzyme responsible has not been well characterised. Two mutants have been described in yeast which affect the SAT reaction (Jones and Fink, 1982). These are known as cys1 and cys2 (see figure 1.3 a). Following isolation and identification of the putative CYS1 gene from a yeast genomic library by mutant complementation (D'Andrea, 1989), it was clear, by lack of nucleotide homology with known bacterial SAT sequences, that the isolated gene does not encode the yeast SAT enzyme. Not surprisingly, the gene product synthesized in vitro has no SAT activity. Rather, the gene has strong homology to E. coli γ-cystathionine synthase, an enzyme which is involved in transulphuration. The cys1 mutation therefore may exert an effect on the SAT reaction directly by inactivating a multienzyme complex involving the SAT enzyme, or indirectly as a metabolic consequence of the block in the interconversion of cysteine and methionine. The pleiotropy seen here added weight to the decision not to pursue the yeast gene for the SAT enzyme as a candidate transgene.

The second reaction step, conversion of OAS to cysteine, is catalysed by a bifunctional enzyme in yeast known as O-acetylserine O-acetylhomoserine sulphydrylase (OAS-OAH sulphydrylase), which can also catalyse the sulphydrylation of O-acetylhomoserine to homocysteine. Both activities are affected in met17 and met25 mutants (presumed to be mutations in separate genes from complementation studies; see fig. 1.3 a; Masselot and de Robichon-Szulmajster, 1975). The MET17 and MET25 genes were isolated independently and when compared, were found to be identical, suggesting that met17 and met25 are in fact mutants in the same structural gene (D'Andrea et al., 1987). In addition, the size and amino acid composition of the protein, as predicted from nucleotide sequence, were consistent with those determined for purified OAS-OAH sulphydrylase (Yamagata, 1976). Given that the enzyme activity attributed to the MET17/25 gene was very well characterised (Yamagata, 1987) this gene seemed a logical choice for further study. Further details of this gene (hereafter

referred to as the MET25 gene for simplicity) and its protein product are given in Chapter 3 of this thesis, where there is also a detailed account of experiments done to determine the suitability of this gene as a transgene for use in the sheep.

1.6.2 The cysteine biosynthesis genes of Salmonella typhimurium.

As in yeast, the pathway to cysteine synthesis in bacteria such as Salmonella typhimurium and Escherichia coli, is a two-step reaction beginning with L-serine and progressing to L-cysteine through the intermediate OAS. This catabolic pathway has been well documented (for a review, see Kredich, 1987). The mutants isolated which define these steps in S. typhimurium, and mutants of the sulphate assimilation pathway which provides the essential substrate sulphide, are shown in figure 1.4. In the cysE mutant, the gene encoding the acetyltransferase enzyme, SAT, is defective, while the cysK and cysM mutants, defective in the second reaction step, represent mutations in genes encoding the OAS sulphydrylase-A and -B enzymes respectively. These enzymes perform the same catalytic function but differ in substrate specificity (in addition to sulphide, the cysM gene product can utilise thiosulphate) and in their physical requirements for activity (the cysM gene product is active as a free dimeric species but the cysK gene product must form a complex with the cysE gene product, SAT, in order to be active). The genes which correspond to these mutations have each been isolated and positively identified (Monroe and Kredich, 1988; Kuczek, Sivaprasad, Bawden and Rogers, manuscript in preparation).

The cysE gene, encoding the SAT enzyme, and the cysM gene encoding the OAS sulphydrylase-B enzyme (thought to be more versatile than OAS sulphydrylase-A, considering the facts given above) were chosen for assessment as transgenes for use in the sheep. Other details related to the cysE and cysM genes and the properties of their protein products will be given in subsequent chapters, along with an account of the experiments done to determine their usefulness as transgenes. The use of the cysK gene in place of cysM is currently under consideration and this will be discussed in the final chapter.

Figure 1.4 : Cysteine biosynthesis in S. typhimurium and E. coli.

The two step conversion of serine to cysteine is shown along with the mutants affecting the acetyltransferase (cysE) and the OAS sulphydrylase (cysM/cysK) reaction steps. In addition, the steps involved in sulphate assimilation are shown, the end-product being sulphide. The diagram is taken from Kredich, 1987.

Sulfate (extracellular)

cysA
cysZ | sulfate permease

Sulfate (intracellular)

cysD | ATP sulfurylase

APS

cysC | APS kinase

PAPS

cysH | PAPS sulfotransferase

Sulfite

cysG
cys I
cys J | sulfite reductase

Acetyl-CoA   Sulfide

cysE   cysM cysK
Serine ──────────► O-Acetylserine ──────────────► Cysteine
      serine   O-acetylserine
   transacetylase   sulfhydrylases A and B

                                                     ▲
                                                     │
Thiosulfate ──────────────────────────► S-Sulfocysteine
                        O-acetylserine
                        sulfhydrylase B

(APS = adenosine 5' phosphosulphate, PAPS = 3' phosphoadenosine 5'
phosphosulphate)

1.7 Availability of substrates for cysteine biosynthesis : the requirement for tissue-specific expression of the transgenes.

As previously discussed, the substrates required for cysteine biosynthesis are L-serine, acetyl-CoA and sulphide. The availability of sulphide in the rumen is the key to the success of the current transgenesis programme. As shown in figure 1.4, a battery of enzymes is involved in sulphate assimilation, the end product being sulphide which is used in the synthesis of cysteine in microorganisms. The enzymes of another group of microorganisms, the sulphur-dissimilatory bacteria contribute to the degradation of sulphur-containing amino acids in the rumen of the sheep (Bray and Till, 1975). This results in the production of large quantities of sulphide (see section 1.4). Following this, the sulphide moves rapidly into the bloodstream by way of the epithelial cells lining the rumen and a large proportion is lost from the sheep by excretion in the urine as sulphate (Bray, 1969a). Consequently, the preferred site for expression of the genes for cysteine biosynthesis is the epithelial lining of the rumen, where this vital substrate could be captured.

Measurements of the amount of sulphide passing out of the rumen of sheep fed a diet of lucerne (given as irreversible loss of sulphide from the rumen; Kennedy and Milligan, 1978) suggest that an average of 1.43 grams of sulphur, as sulphide, moves from the rumen each day. In that work, it was proposed that greater than 90 % of this loss from the rumen may be accounted for by absorption, as only small amounts of sulphide flow from the rumen in digesta. This means that an average of at least 1.29 grams of sulphur, as sulphide, traverses the epithelial lining of the rumen each day. The aim of the current transgenesis programme is the capture of this sulphide by the OAS sulphydrylase enzyme. Utilisation of this level of sulphide would enable the synthesis of 4.88 grams of cysteine per day. Given that 100 % capture and conversion could not be expected, but also that as little as 0.5 grams / day of abomasally infused cysteine stimulates wool growth (Reis, 1967) an improvement in wool growth rate could reasonably be expected if little more than 10 % of the sulphide was captured in transit through the epithelial lining of the rumen. Assuming an efficient and

similarly rapid flux of newly synthesized cysteine from the epithelial cells into the bloodstream, the potential for improved wool growth would exist.

In order to achieve rumen-specific expression, the promoters of genes expressed at a high level in this tissue are being sought. A number of rumen cDNA clones are currently under investigation, including several rumen keratin cDNAs (L.A. Whitbread, Ph.D. thesis) and some non-keratin rumen cDNA clones (C.J. McLaughlan, pers. comm.). The isolation of the genes and their promoters is beyond the scope of this thesis, hence expression of the MET25, cysE and cysM genes will be discussed in the context of the other eukaryotic promoters to which they have been linked during this work. In relation to the function of a cysteine biosynthesis pathway in the rumen epithelium and the consequent utilisation of sulphur, consideration of the normal mechanisms by which the sheep maintains a sulphur balance in its body must be given. This will be discussed further in the final chapter.

1.8   Transfer of the cysteine biosynthesis genes to the sheep genome.

The transfer of any of the cysteine biosynthesis genes to the sheep genome must be preceded by a rigorous assessment of their suitability for expression in a mammalian environment (see 1.9). The genes which function effectively in mammalian systems to produce the catalytic capacity of the authentic yeast or bacterial enzymes, will be used in the sheep transgenesis programme. As the method of trangenesis in mammals, and in particular livestock, has been well documented, it will not feature as a major discussion point in this thesis. (For reviews of the traditional methods and the more recent additions to these, for example homologous recombination in embryonic stem cells, see Jaenisch, 1988; Pursel et al., 1989; Capecchi, 1989). Improvements made to the techniques of producing viable lambs from microinjected ova (Walker, 1990; Walker et al., 1990) will only be mentioned in the context of the higher rate of transgenesis achieved. The options which exist for sheep trangenesis with these genes subsequent to the work presented here, will be discussed in the final chapter. Also, a feature central to the transgenesis attempted here is that the two required genes must be supplied to the sheep genome so that they may be coordinately

expressed, and also inherited as a single, bifunctional unit of the genome. This will be discussed fully in Chapter 5.

1.9 The research aims embodied in this thesis.

As introduced in the preceding pages, the overall aim of the work for this thesis was to take steps towards the establishment of a heritable trait of increased wool growth through transgenesis of the sheep with the microbial genes for cysteine synthesis. In essence, this would result in the provision of the sheep with a new amino acid synthesis pathway, a cysteine biosynthesis pathway which is not normally found in mammals. Before the sheep can be provided with this two-step enzyme pathway, certain constituent aims must be met. These involve analysis of the cysteine biosynthesis genes, one from the yeast S. cerevisiae, the MET25 gene, and two genes from the bacterium S. typhimurium, the cysE and cysM genes, with respect to the fidelity of their expression in a mammalian environment, when specifically constructed for eukaryotic expression. This includes assessment of the enzyme activity of the products of these genes when they are expressed in vitro, in the rabbit reticulocyte lysate and in cell culture and in vivo in transgenic mice. A thorough investigation of the behaviour of these genes in the mammalian systems mentioned will precede transgenesis of the sheep with similar gene constructs in which the expression is tissue-specific and rigidly controlled by the inclusion of appropriate processing signals.

Chapter 2

Materials and Methods.

Chapter 2 : Materials and Methods.

2.1 Materials.

To follow is a list of the materials used in the course of the experimental work described in this thesis. Most are general reagents; the source of those specific to a particular method is mentioned in the Methods section.

2.1.1 Chemicals.

The chemicals used are listed below, along with their source. All chemicals used were analytical grade, or of the highest available purity.

Sigma Chemical Company.: acrylamide and bis acrylamide (N,N'-methylene-bis-acrylamide), ampicillin, adenosine triphosphate, arginine (free base), BCIG (5-bromo-4-chloro-3-indolyl3-D-β-galactoside), deoxyribonucleoside triphosphates, DTT (dithiothreitol), DTE (dithioerythritol), ethidium bromide, EDTA (ethylenediaminetetracetic acid), Haemin, IPTG (isopropylthiogalactoside), B-mercaptoethanol, phenylmethanesulphonyl fluoride (PMSF), PIPES [1,4-piperazinebis(ethanesulphonic acid)], polyvinyl pyrrolidone, potassuium chloride, potassium acetate, pyridoxal 5'-phosphate, salmon sperm DNA, L-serine, sodium acetate, SDS (sodium dodecyl sulphate), sarkosyl (sodium dodecyl sarcosinate), spermidine, spermine, TEMED (N,N,N,N'-tetramethylethylenediamine), Tween 20, yeast tRNA, zinc chloride

Merck : guanidinium isothiocyanate, HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid], tetra-sodium pyrophosphate, sodium salicylate, Tris [Tris(hydroxymethyl)aminomethane], Proteinase K

Boehringer Mannheim : Bovine Serum Albumin (BSA; nuclease free)

BDH : Bromophenol blue, Butan-1-ol, calcium chloride, Dimethyl sulphoxide (DMSO) Hydrochloric acid, Formamide, Isopropanol, magnesium chloride, magnesium acetate, Nonidet-P40 (NP40).polyethlene glycol(8000) sucrose, urea.

Bresatec Pty. Ltd. : NBT and BCIP.

May and Baker : Boric acid, glacial acetic acid, sodium chloride, sodium citrate sodium hydroxide

Pharmacia : Dextran sulphate, ficoll 400, Sephadex G-50 (Superfine)

BRL : Agarose (type 1) and Low melting temperature agarose

Ajax Chemical Co. : Ammonium persulphate, ammonium sulphate, Chloroform, D-glucose, Isoamyl alcohol, magnesium sulphate, potassium dihydrogen phosphate, potassium hydrogen phosphate, potassium hydroxide, sodium sulphide

Wako Pure Chemical Industries Ltd. : Phenol (redistilled)

Tokyo Kasei : Xylene cyanol.

Cabot : Caesium chloride (technical grade)

F.S.E. Scientific : ethanol (supplied as a 95% solution, then redistilled)

Commonwealth Industrial Gases (CIG) Pty. Ltd.: liquid nitrogen and carbon dioxide.

* Acetyl-CoA was a gift from G Booker and C. Hahn.

* 4-vinylpyridine was a gift from Dr. A. Inglis (CSIRO Division of Protein Chemistry, Parkville, Melbourne) and was redistilled before use.

2.1.2   Enzymes.

Restriction Endonucleases were purchased from either New England Biolabs, Pharmacia or TOYOBO Co. Ltd..

Bal31 Exonuclease was obtained from Promega.

E. coli DNA Polymerase I and the Klenow fragment, Deoxyribonuclease I, T4 bacteriophage DNA Ligase, T4 Polynucleotide Kinase and bacteriophage SP6 and T7 RNA Polymerases were purchased from Bresatec Pty. Ltd.

Lysozyme, Ribonuclease T1 and Ribonuclease A were purchased from Sigma Chemical Co.

Calf Intestinal Phosphatase was purchased from Boehringer-Mannheim.

Proteinase K was purchased from Merck.

"AmpiTaq" Taq DNA polymerase was purchased from Cetus (per Bresatec)

AMV Reverse transcriptase was purchased from Pharmacia.

2.1.3   Bacterial Strains.

The salient features of the genotypes of the E. coli K12 bacterial strains used in the recombinant DNA work are given.

MC1061 : ara D139,. D(araABC-leu) 7679, DlacX74, gal U$^-$, gal K$^-$, hsr$^-$, hsm$^+$, str A (Casdaban and Cohen, 1980). Host for transformation with pBR322-based vectors.

ED8799 : hsdS, metB7, supE (glnV) 44,supF (tyrT) 58, lacZDM15, r_k- m_k- (a gift from Dr. S Clarke, Biotechnology Australia). Host for transformation with pBR322-based vectors.

JM101 : supE, thi D(lac-pro AB) F' [traD36 proAB$^+$ lacI^q ZDM15] (Messing, 1979). Host for M13 transfections.

2.1.4   Plasmids and Bacteriophage strains

Plasmids used as base vectors in cloning and for the isolation of specific sequences are listed below.

pIBI76 : purchased from International Biotechnologies, Inc.

pGEM-1 and pGEM-2 : purchased from Promega.

pSP64-T : a gift from P Krieg.

pJL4 : (Gough et al., 1985) a gift from F Shannon.

pJL-E : a gift from E. Kuczek.

pRSVΔH (pRSVN.03) : a gift from A Robins.

pLT-E : a gift from E. Kuczek.

pCS942 : a gift from A Sivaprasad.

Bacteriophage M13 mp18 and mp19 : (Messing, 1983) were purchased from Bresatec Pty. Ltd.

2.1.5   Bacterial Growth Media.

All solutions and media were prepared using water purified using the Millipore Reverse Osmosis water purification system (double-distilled equivalent) and sterilized by autoclaving.

Luria Broth : 1 % Bactotryptone (Difco) 0.5 % yeast extract (Difco) 1.0 % NaCl, pH 7.2; used for the growth of E. coli MC1061 and ED8799. Agar plates were prepared by adding 1.5 % Bacto-agar (Difco) to the above media. Where appropriate, the media and plates were supplemented with ampicillin (Sigma) at 50 μg/ml.

M13 minimal medium : 1.05 % potassium hydrogen phosphate, 0.45 % potassium dihydrogen phosphate 0.1 % ammonium sulphate, 0.05 % sodium citrate and after autoclaving, was supplemented with 0.02 % magnesium sulphate, 0.0005 % Thiamine-

HCl and 0.2 % glucose; used for the growth of E. coli JM101. M13 minimal plates were prepared by addition of 1.5 % Bacto-agar.

2xYT Broth : 1.6 % Bacto-tryptone, 1 % yeast extract, 0.5 % NaCl, pH 7.0; used for the growth of E. coli JM101 infected with M13. Soft YT agar overlay was prepared by the addition of 0.7 % Bacto-agar.

2.1.6 Buffers and Solutions

General Buffers

TE : 10 mM Tris-HCl (pH 7.5 or 8.0), o.1 mM EDTA.

SSC : 150 mM NaCl, 20 mM sodium citrate, 1.0 mM EDTA, pH 7.0.

SSPE : 150 mM NaCl, 20 mM sodium dihydrogen phosphate, 1.0 mM EDTA, pH 7.4.
Tris-PLP : 10 mM Tris-HCl (7.5), 0.1 mM pyridoxal 5' phosphate

Tris-PLP/PMSF : 10 mM Tris-HCl (7.5), 0.1 mM pyridoxal 5' phosphate, 1.0 mM PMSF.
Electrophoresis Buffers.

1 x TAE : 40 mM Tris acetate, 20 mM sodium acetate, 1 mM EDTA, pH 8.2

1 x TBE : 90 mM Tris, 90 mM Boric acid, 2.5 mM EDTA, pH 8.3.

1 x Tris-glycine-SDS : 25 mM Tris, 0.192 M glycine, 1.0 % SDS, pH 8.3

3 x Agarose Loading Buffer : 50 mM Tris-HCl (pH 7.5) 50 % glycerol (v/v) 5 mM EDTA, 0.1 % bromophenol blue and 0.1 % sarkosyl.

2 x DNA Formamide Loading Buffer : 95 % formamide (v/v),10 mM NaOH, 20 mM EDTA, 0.03 % xylene cyanol and 0.03 % bromophenol blue

1 x RNA Formamide Loading Buffer : 95 % formamide (v/v), 20 mM EDTA, 0.06% xylene cyanol and 0.06% bromophenol blue

2 x Protein Loading Buffer : 10 % glycerol (v/v) 2.5 % SDS, 25 mM Tris (pH 8.3) and 5.0 % B-mercaptoethanol (v/v).

Prehybridisation/Hybridisation Solutions

Prehybridisation/Hybridisation solution for Genescreen/Zetaprobe membranes (Southern transfer analysis) : 0.2 % BSA, 0.2% polyvinyl-pyrrolidone (40,000) 0.2 % ficoll (400,000) 50 mM Tris-HCl (pH 7.5) 0.1 % sodium pyrophosphate, 1.0 % SDS, 10 % dextran sulphate (500,000), 50 % formamide, 1.0 M NaCl and 500 µg/ml salmon sperm DNA (random sheared and heat denatured).

Prehybridisation/Hybridisation/Wash solutions for Nitrocellulose membrane (Western transfer analysis):

Buffer 1 : 1 x PBS-Tween 20, 1 % BSA.

Buffer 2 : 1 x PBS-Tween 20.

Buffer 3 : 100 mM Tris-HCl (pH 9.5) 100 mM NaCl, 5 mM magnesium chloride.

Buffer 4 : 10 mM Tris-HCl (pH 7.5) 1 mM EDTA.

* 1 x PBS-Tween 20 is 1 % Tween 20 in PBS (Phosphate buffered saline).

2.1.7   Radiochemicals

DNA labelling : [α-^{32}P] dATP and [α-^{32}P] dCTP (specific activity 3000 Ci/mmole) were purchased from Bresatec Pty. Ltd.

RNA labelling : [α-^{32}P] rUTP (specific activity 3000 Ci/mmole) was purchased from Bresatec Pty. Ltd.

Protein labelling : L-[3-^{3}H] serine (30 Ci/mmole) and L-[^{35}S] methionine (>800 Ci/mmole) were purchased from Amersham Australia Pty. Ltd..

Enzyme assays : L-[U-^{14}C] serine (171 mCi/mmole) was purchased from Amersham Australia Pty. Ltd..

2.1.8   Molecular Biology Kits.

Kits used for the preparation of DNA probes by Nick translation, synthesis of SP6 and T7 sense (non-radioactive) and anti-sense (radiolabelled) transcripts and for the sequencing of single-stranded M13 DNA were purchased from Bresatec Pty. Ltd..

The Geneclean DNA purification kit (BIO 101) was also purchased from Bresatec Pty. Ltd.

Rabbit reticulocyte lysate and materials for the in vitro translation reactions were purchased from Amersham Australia Pty. Ltd.

2.1.9   Equipment.

Specific equipment used in the experimental work is detailed in the Methods sections in 2.2. Other than this, all plasticware used in bacterial work was sterile, and all glassware used in this and other work were sterilised by autoclaving prior to use. In addition, all micromanipulations of solutions were carried out with Gilson "Pipetman" pipetting aids.

2.1.10   Cell Culture Materials.

Listed below are the sources of reagents and equipment used in the cell culture work. Materials not mentioned here are referred to more specifically in the Methods section 2.2.4. All heat-labile solutions and media were filter sterilised before use.

Foetal calf serum (FCS) : Commonwealth Serum Laboratories (CSL).

Trypsin (Difco) was made as a 0.1 % solution in 1 x Versene (EDTA; CSL).

Dulbecco's Modified Eagle's Medium used in the culture of HeLa, COS-1, RESV40 and CSL503 cells was purchased from Gibco

Ham's F12 Medium used in the culture of CHO cells was purchased from Gibco.

Phosphate buffered saline (PBS) : 136 mM NaCl, 2.6 mM potassium chloride, 1.5 mM potassium phosphate, 8.0 mM sodium phosphate, pH 7.3.

HEPES buffered saline : 20mM HEPES, 137 mM NaCl and 5 mM potassium chloride, pH 7.05. (HEPES buffered saline with 6 mM glucose was used as the CSL503 electroporation buffer)

DEAE-Dextran used in the HeLa cell transfections was purchased from Pharmacia.

Plasticware (sterile) used was obtained from either Linbro (multiwell trays), Lux (Nunc. Inc.) (60 and 100 mm dishes), Terumo (syringes), Sartorius (micro-filtration cartridges) and Disposable Products (10 and 20 ml polypropylene tubes).

"Pipet-aid" pipetting devices (Drummond Scientific Co.) were used in work carried out in the Biohazard and Laminar flow hoods.

2.1.11   Animals used in the Transgenesis programmes.

The mice used in transgenesis were of the C57BL/6J mouse strain and the sheep were of the South Australian Merino breed. All animals were maintained and handled according to the guidelines specified by the Genetic Manipulation Advisory Committee (Australia) and enforced by the Biohazards and Animal Ethics Committees of the University of Adelaide.

2.2   Methods.

To follow is a list of the methods used in the experimental work undertaken in this thesis, distinguished as DNA, RNA, protein, cell culture and transgenesis methods (sections 2.2.1 - 2.2.5 respectively). Detail of a particular method is only given where it

departs significantly from the original method described, or if it is a new method devised for use in this work. Many of the methods used were adapted from those described in the "Molecular Cloning" manuals of Maniatis et al. (1982) and Sambrook et al. (1989) and will be referenced accordingly. Similarly, the transgenesis methods used in this work are described in "Manipulating the Mouse Embryo" by Hogan et al. (1986). In the case of protocols not described in these texts, reference to original method(s) used will be given. Finally, the computer programmes used in the analysis of DNA sequence, RNA structure and in quantitative densitometric analysis of autoradiograms are referenced in section 2.2.6.

2.2.1 DNA Methods.

2.2.1.1 Recombinant DNA methods.

The following listed methods were performed essentially as described by Maniatis et al. (1982) and Sambrook et al. (1989) unless indicated otherwise, in which case the appropriate reference is given. Supplementary information is given with some of the methods. Handling of recombinant DNA in viable organisms (C1) and of recombinant DNA in non-viable organisms (C0) was carried out according to the rules of containment specified by the Australian Academy of Science Committee on Recombinant DNA and the University Council of the University of Adelaide.

Enzymatic modification of DNA.

(a) Restriction endonuclease digestion.

(b) Bal 31 exonuclease digestion.

(c) End-filling of 5' and 3' protruding termini using the Klenow fragment from E. coli DNA Polymerase I.

(d) Dephosphorylation of 5' and 3' phosphoryl termini (protruding or blunt ends) using Calf Intestinal Phosphatase.

Purification of DNA following enzymatic modification.

(a) Phenol extraction.

(b) Chloroform : isoamyl alcohol extraction.

(c) Ethanol precipitation.

(d) Spermine precipitation : as described by Davis et al. (1986).

(e) Fractionation via (TAE) agarose or (TBE) polyacrylamide gel electrophoresis (see general DNA methods).

Recovery of DNA fragments following fractionation by gel electrophoresis.

(a) Extraction from agarose by the method described by Vogelstein and Gillespie (1979) using a "Gene-clean" kit obtained from BIO 101, used according to the manufacturers instructions.

(b) Extraction from low melting-point 1 % (TAE) agarose.

(c) Adsorption to and elution from NA45 ion-exchange paper (Schleicher and Schuell) using the method described by Ken Burtis (pers. comm.). Briefly, DNA electrophoresed in a horizontal agarose gel in TAE containing ethidium bromide is electrophoretically transferred to a thin strip of NA45 paper (inserted into a scalpel cut made in the path of the band migrating in the gel) at 100 mA. After the transfer is complete (can be visualised under long wave U.V. light) and the paper rinsed in T.E. (pH 8.0) the DNA may be eluted from the paper in 0.4 ml of 1.0 M NaCl/0.05 M Arginine (free base) in an eppendorf tube at 70 ^{0}C. After 1-3 hours, DNA in the eluate may be precipitated with 2.5 volumes of ethanol.

(d) Elution from 1 % (TAE) agarose or 5 % (20 : 1 acrylamide : bis-acrylamide) non-denaturing acrylamide.

Ligation of vector and insert DNA molecules. : Ligation reactions were performed overnight at 14 ^{0}C in a 20 µl volume using 1.0 unit of T4 bacteriophage DNA Ligase, 10 - 50 ng of vector DNA and enough insert DNA to provide a minimum 3 : 1 and maximum 1 : 3 mole vector : insert ratio (with respect to the insert) depending upon the size of the insert fragment. The ligation buffer used was as described in Maniatis et al. (1982).

Transformation of E. coli. strains. : Fresh competent cells (MC1061, ED8799 and JM101) were prepared and transformed using the calcium chloride method described by Dagert and Ehrlich (1979). Half of a ligation reaction was routinely used in the transformation and usually provided a sufficient number of recombinants for analysis. Recombinant clones were selected using nutrient agar plates supplemented with appropriate antibiotics or inducers and chromogenic substrates, plasmid or replicative form DNA prepared (see below) and this

analysed by restriction endonuclease digestion. M13 clones were subsequently analysed by dideoxy DNA sequence analysis (see below).

Preparation of plasmid DNA.

(a) Minipreparations (2 ml cultures) of plasmid recombinants using the alkaline lysis procedure originally described by Birnboim and Doly (1979) and Ish-Horowitz and Burke (1981).

(b) Large scale preparations (50 ml cultures) of plasmid clones using the alkaline lysis procedure.

(c) CsCl equilibrium density gradient purification of plasmid DNA from large scale preparations. Continuous gradients were used and generated by centrifugation of 2.2 ml samples in the Beckmann TL-100 Ultracentrifuge for 16 hours at 80,000 revolutions per minute.

Methods of analysis of DNA sequence.

(a) Preparation of single-stranded DNA from M13 clones (Messing, 1983).

(b) Nucleotide sequence determination by the dideoxy chain-termination method of Sanger et al. (1977).

(c) Analysis of multiple Bal 31 deletion M13 clones via the use of single dideoxy-T reactions.

(d) Preparation of Replicative Form DNA from M13 clones.

2.2.1.2   General DNA methods.

Analysis of DNA by gel electrophoresis.

(a) Analytical and preparative agarose gel electrophoresis : carried out in horizontal apparati as described by Maniatis et al. (1982) in 1 % or 2 % (TAE) agarose gels. Small gels approximately 55 mm x 75 mm in size were routinely used for analysis of restriction endonuclease digestions of plasmid or genomic DNA and also on a preparative scale for the purification of DNA insert and probe fragments ranging in size from 300 - 4000 bp. Larger horizontal gel apparati (Pharmacia) were used in the fractionation of DNA prior to Southern transfer analyses. Electrophoresis in the small gel apparati was carried out at a constant current of 100 mA and at 150 - 250 mA in the larger units. Samples were

resuspended in agarose DNA loading buffer (see Materials) prior to gel electrophoresis.

* Sizes of DNA fragments in analytical and preparative gel electrophoresis were determined by comparison of the migration distance of the unknowns with those of marker DNA fragments in HpaII-digested pUC19 DNA, EcoRI-digested bacteriophage SPP-1 DNA and HindIII-digested lambda phage DNA co-electrophoresed with the test DNA samples. The pre-digested marker DNA samples were purchased from Bresatec Pty. Ltd..

(b) Analytical and preparative polyacrylamide gel electrophoresis : carried out in vertical apparati as described by Maniatis et al. (1982) in 5 % (20 : 1 acrylamide : bis-acrylamide) non-denaturing or denaturing gels (both with TBE buffer) approximately 140 mm x 120 mm in size. Denaturing gels contained 7 M urea in addition to the other components. These gels were used to fractionate mixtures of small DNA fragments (ranging in size from 50 - 300 bp) for analytical or preparative purposes. Pre-electrophoresis of the gels was conducted for 20 minutes at a constant current of 20 mA, and this current resumed after loading of the gel.

Samples to be loaded onto the denaturing gels were usually dissolved in DNA formamide loading buffer (see Materials) and heated at 100 ^{0}C for 3 - 5 minutes prior to loading.

Preparation of genomic DNA.

(a) Isolation from the nuclei of cultured cells (following NP40 lysis to obtain cytoplasmic RNA; see section 2.2.2)

(b) Isolation from mouse and sheep tissues.

* Crude cell and tissue extracts were treated first with 150 μg/ml Proteinase K for 3 - 16 hours at 37 ^{0}C and high molecular weight DNA purified by phenol/chloroform extraction as described by Maniatis et al. (1982).

Determination of DNA concentration. : Performed by measurement of absorbance of 260 nm wavelength U.V. light (Shimadzu UV-visible recording spectrophotometer, model UV 160A) where a solution containing 50 μg/ml of double-stranded DNA has an absorbance = 1.0.

Southern transfer. : Transfer of restriction endonuclease digested genomic DNA to Zetaprobe membrane was performed in 0.4 M NaOH using either a

modification of the capillary blotting method of Southern (1975) described by
Reed and Mann (1985) or by vacuum blotting using the LKB "Vacugene"
manifold according to the manufacturers instructions.

DNA Dot blot. : Transfer of undigested, heat-denatured genomic DNA by
vacuum blotting in 0.1 M NaOH to Genescreen membrane was carried out using
a Schleicher and Schuell "MINIFOLD" vacublotter according to the
manufacturers instructions. Copy control series dotted alongside test samples
were made by serial dilution of plasmid DNA and combination of this with an
amount of negative control genomic DNA equivalent to that being analysed in the
test samples. After neutralisation of the blotted samples with 0.5 M Tris.HCl (pH
7.5) the transfer membrane was baked for 2 - 4 hours at 80 ^{0}C under vacuum.

Hybridisation of radio-labelled DNA probes to membrane-bound DNA. : Radio-
labelled DNA probes prepared by Nick-translation (see below) using ^{32}P-dNTP's
(Bresatec)were hybridised to DNA bound to Zetaprobe and Genescreen
membranes for 12 - 16 hours at 42 ^{0}C in the presence of formamide, dextran
sulphate, Denhardt's reagent and 500 µg/ml heat-denatured, sheared salmon
sperm DNA (see Materials section 2.1.6). (Prehybridisation of transfer
membranes was carried out for 12 - 16 hours at 42 ^{0}C using the same solution;
heat denatured probe DNA was usually added directly to this to begin the
hybridisation step.)

Post-hybridisation washing of membranes. : Removal of non-specifically bound
probe molecules was attempted first by washing the transfer membrane under low
stringency conditions (twice in 2 x SSC / 0.5 % SDS at room temperature for 5
minutes with agitation, then twice more in the same solution for one hour at 65 ^{0}C
with occasional agitation). At this stage, autoradiography was carried out for 2 - 4
hours at -80 ^{0}C using pre-flashed film. If necessary, the membrane was rewashed
at a higher stringency (usually at 0.2 x SSC / 0.5 % SDS at 65 ^{0}C for at least one
hour) and autoradiography repeated using normal film, in this case for 12 - 16
hours.

Radio-labelling of DNA.

(a) Nick-translation of DNA probes : performed in a 20 µl volume using
100 - 200 ng of purified linearised probe DNA, 50 µCi of α-^{32}P-dCTP and 25
µCi of α-^{32}P-dATP, 1 x buffer without dCTP, with only 2.0 µM dATP and with

200 µM dGTP and dTTP, and an enzyme mixture (pancreatic DNase I and E. coli DNA Polymerase I) prepared by Bresatec. After 90 minutes at 14 ^{0}C, reactions were terminated by the addition of 2.0 µl of 0.5 M EDTA (pH 8.0). Radio-labelled DNA was then separated from unincorporated radio-nucleotides by size-fractionation using a G50 Sephadex ("Superfine") column, with TE (pH 8.0) / 5 mM B-mercaptoethanol as the running buffer. Probe DNA collected in 200 - 400 µl was heat denatured at 100 ^{0}C for 5 minutes, combined with 1 - 2 ml of fresh prehybridisation solution, then added to the prehybridised membrane to begin the hybridisation.

(b) End-labelling of 5' protruding termini : performed in a 20 µl volume at 37 ^{0}C using the Klenow fragment of E. coli DNA Polymerase I, Nick-translation buffer (without dCTP) and 20 µCi of α-^{32}P-dCTP. After 30 minutes, reactions were terminated by addition of 2.0 µl of 0.5 M EDTA (pH 8.0). TE (pH 8.0) was then added to make the volume 200 µl and the mixture extracted with phenol/chloroform/isoamyl alcohol. Following this, the aqueous phase was stored at -20 ^{0}C prior to use. This method was routinely used to prepare labelled HpaII-digested pUC19 DNA fragments for use as size markers.

(c) End-labelling ("Kinasing") of the 5' termini of synthetic oligonuleotide primers : performed in a 10 µl volume using 50-100 ng of primer DNA, approximately 1.0 unit of T4 bacteriophage Polynucleotide kinase, 50 µCi of γ-^{32}P-ATP and in the buffer described by Maniatis et al. (1982). After 30 minutes at 37 ^{0}C, reactions were terminated by addition of 5 µl of DNA formamide loading buffer and radio-labelled primers purified by fractionation in a 20 % non-denaturing polyacrylamide (30 : 1 acrylamide : bis-acrylamide) gel (200 x 200 x 0.5 mm). Samples were heated at 100 ^{0}C for 3 minutes prior to loading onto the gel. After autoradiographic detection in situ, primers were eluted from excised gel slices in 400 µl TE (pH 8.0) at 37 ^{0}C for 4-16 hours. Eluted primers were then stored at -20 ^{0}C until use in primer extension analysis.

Polymerase chain reaction. : PCR reactions were set up using components of a Perkin-Elmer Cetus "GeneAmp" DNA Amplification kit and in the way described by the manufacturer, according to the method of Saiki et al. (1988). Reaction cycles were performed in the Perkin-Elmer Cetus DNA Thermal Cycler, using a 40 cycle programme consisting of consecutive 94 ^{0}C / 30 second denaturation, 55

^{0}C / 15 second annealing and 72 ^{0}C / 30 second extension steps. The first cycle was preceded by a 95 ^{0}C / 10 minute denaturation step. The enzyme used in the kit was the thermostable recombinant DNA Polymerase "AmpliTaq" derived from a modified form of the gene from Thermus aquaticus. Primers used were derived from the coding region of the cysM gene of S. typhimurium and the sequence of each is shown below.

Primer A : 5'-GCTACAGGACTAACTTCGCTGGTCG-3'

Primer B : 5'-GCATTGCCGCCCGACGCTCCTGACTC-3'

Products from the PCR reactions were fractionated in 5 % non-denaturing polyacrylamide (20 : 1 acrylamide : bis-acrylamide) gels and visualised by short wave U.V. after staining with ethidium bromide.

2.2.2 RNA Methods.

Unless another reference is given, the RNA methods listed below were carried out essentially as described in Maniatis et al. (1982) and Sambrook et al. (1989).

Isolation and purification of mRNA.

(a) NP40 lysis method to obtain cytoplasmic RNA.

(b) Acid-guanidine thiocyanate-phenol-chloroform method : performed as described by Chomczsynski and Sacchi (1987) to isolate total cellular RNA from sheep tissues.

(c) Determination of RNA concentration : by absorbance of 260 nm wavelength U.V. light, standardised with respect to an absorbance of 1.0 for a 40 µg/ml solution of single-stranded DNA at 260 nm.

(d) Analysis of RNA by gel electrophoresis : Rapid analysis of the integrity of RNA preparations was performed using gel electrophoresis in 1 % (TAE) agarose . Aliquots of aqueous RNA samples were loaded in non-denaturing agarose DNA loading buffer onto small analytical gels, electrophoresed at 100 mA then visualised with short wave U.V. light following staining with ethidium bromide. RNase free conditions were maintained during preparation of the gel and gel buffer, and gel tanks were copiously rinsed in water purified in the Milli-RO (Millipore) system.

<u>Synthesis and purification of in vitro transcripts.</u>

(a) Non-radioactive SP6 and T7 sense transcripts (for RRL translations and as positive control transcripts in the RNA protection assays) : performed according to the method of Melton et al. (1984). Transcripts intended for RRL translations were synthesised with addition of the 5'm7G analogue m7G(5')ppp(5')G according to the method of Krieg and Melton (1987) and using reagents and enzymes supplied with an SP6/T7 transcription kit obtained from Bresatec. Post-synthesis processing included increasing the volume to 200 μl with TE (pH 8.0), a phenol extraction then a back extraction with 100 μl of additional TE (pH 8.0), a separate chloroform / isoamyl alcohol extraction of the 300 μl aqueous phase then ethanol precipitation using 0.1 volumes of 4.0 M NaCl and 2.5 volumes of ethanol at -20 ^{0}C. RNA pelleted by centrifugation at 11,600 x g in a microfuge (Microcentaur, MSE) was washed in sterile redistilled ethanol, dried under vacuum (Speed Vac Concentrator, Savant) resuspended at approximately 0.5 μg/μl in RNase free H_2O and stored at -20 ^{0}C until use. A small aliquot was usually examined by electrophoresis in a 1 % TAE agarose gel as described above.

SP6 transcripts intended for RRL translations were trace-labelled by addition of α-^{32}P-rUTP to transcription reactions. The RNA concentration was then calculated according to the percentage incorporation of the radio-labelled rUTP, as judged by Cerenkow counting (LKB Scintillation Counter) of ^{32}P present in the pellet and supernatant from the ethanol precipitation step.

(b) Radio-labelled SP6 antisense transcripts : synthesised according to the method of Melton et al. (1984) using α-^{32}P-rUTP and reagents and enzymes from an SP6/T7 transcription kit obtained from Bresatec. Probe RNAs were purified by fractionation in preparative 5 % denaturing polyacrylamide gels (composition as used for DNA sequencing). Samples were resuspended in RNA formamide loading buffer (see Materials) prior to gel electrophoresis. Excision of gel slices containing radio-labelled transcript bands after autoradiography preceded elution in 400 μl TE / 0.1 % SDS at 37 ^{0}C for 2 - 3 hours. Following this, the eluate was combined with 0.1 volumes of 4.0 M NaCl and 2.5 volumes of ethanol and stored at -20 ^{0}C for up to two weeks prior to use.

<u>Detection of specific RNA transcripts.</u>

(a) RNA protection analysis : performed as described by Krieg and Melton (1987). Following the hybridisation and post-hybridisation processing steps, half of each protection assay sample was fractionated using 5 % denaturing polyacrylamide (as used for DNA sequencing) and products detected by autoradiography at -80 ^{0}C (with the use of an intensification screen). Samples were resuspended in RNA formamide loading buffer prior to gel electrophoresis. Positive control protection assays were created by the combination of known quantities (e.g. 50 - 500 pg) of synthetic non-radioactive sense transcripts (derived from pGEM clones of specific genes created for RRL translations) with an amount of yeast tRNA equivalent to that being analysed in the test samples. These were then treated as per the test samples throughout the procedure. Two "probe control" reactions, Y- and Y+, consisted of a combination of yeast tRNA and probe transcript (used in amounts comparable to test reactions). One of these (Y-) was not treated with the RNase enzymes but was processed through all remaining steps along with the other samples (including Y+). The Y- control served to show the state of the antisense RNA probe molecules at the conclusion of the procedure and the Y+ control to demonstrate the effectiveness of the RNase digestion treatment in eliminating all non-specific RNA / RNA hybrids at the conclusion of the hybridisation step.

(b) Primer extension analysis : conducted according to the method described by McKnight <u>et al.</u> (1981). Ethanol precipitated reaction products were resuspended in DNA formamide loading buffer, heated at 100 ^{0}C for 5 minutes then fractionated by electrophoresis at 30 mA (constant current) in denaturing 5 % polyacrylamide gels (200 x 400 x 0.5 mm). Prior to autoradiography, gels were treated by gradual washing with 1 litre of 12 % acetic acid (over 15 minutes) and with 1.5 litres of 20 % ethanol (over 15 minutes) to fix nucleic acids in the gel, then drying at 110 ^{0}C for 30 minutes. MET25 and cysE transcription in HeLa cells was analysed in this way, and the sequences of the primers used are given below.

MET25 primer : 5'-GTGGTCCCAAGAGAACCGGCCGCAC-3'
cysE primer : 5'-AAGTGTGCCGTACCGAATGAGTG-3'

2.2.3 Protein Methods.

In vitro translation of proteins in the rabbit reticulocyte lysate (RRL) : Synthesis of proteins in the RRL was performed using the Amersham RRL Translation kit according to the instructions provided, and where proteins were to be radio-labelled to facilitate detection, either L-^{35}S-methionine or L-^{3}H-serine were used. SP6 transcripts were protected from degradation during the translation reactions by the addition of the RNase inhibitor Rnasin (Promega). Quantitative translations were performed using trace-labelled SP6 synthetic transcripts derived from the MET25, cysE and CYS1 pGEM clones described in Chapter 3. Each 15 μl reaction contained 10 μl of the RRL, 5 μCi L-^{3}H-serine (Amersham), 1.5 μl of a 1.0 mm amino acid mixture minus serine (Amersham) 0.5 μl (20 units) of RNasin, 1.5 μl of 1.2 M potassium acetate and 1.5 μl of H$_2$O containing the desired amount of SP6 RNA. Reaction mixtures were incubated at 30 ^{0}C for 90 minutes and one third (5 μl) of each sample electrophoresed in an SDS-polyacrylamide gel (see below). Prior to gel loading, an equal volume of a protein loading buffer (see Materials) was added to each 5 μl sample aliquot and the mixture heated at 80-100 ^{0}C for 5 minutes.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of proteins. : Fractionation of proteins using a continuous pH (8.3) Tris-glycine SDS-PAGE system (Laemmli, 1970) was performed using a 4 % stacking gel (6.5 ml) poured above a 12.5 % separating gel (15.0 ml) to make a gel of dimensions 140 x 140 x 1.1 mm. Proteins were electrophoresed through the stacking gel at 40 mA, then at 25 mA through the separating gel until the tracker dye had migrated to within 10 mm of the base of the separating gel.

Fluorography following SDS-PAGE. : At the conclusion of the SDS-PAGE run, gels were soaked for 15 minutes in 7 % acetic acid with agitation, then transferred to a 1.0 M sodium salicylate solution and soaked for a further 15 minutes with agitation. The gels were removed from this solution and dried under vacuum at 65 ^{0}C onto Whatman 3MM paper using a Hoeffer gel drying apparatus. Autoradiography was then conducted at -80 ^{0}C using an intensification screen (Du Pont) and where quantitation was required, the X-ray film used was pre-flashed.

Preparation of protein extracts.

(a) Cell culture protein extracts : Cells in monolayer cultures were harvested by trypsinisation, washed twice with 5 ml PBS, then resuspended in 500 µl of Tris-PLP buffer (see Materials) in a 1.5 ml microfuge tube and washed three times in this, with gentle pelleting of the cells between washing steps. After the final wash, the cells were resuspended in 100-200 µl of the Tris-PLP/PMSF buffer (see Materials) and lysed by three successive freeze-thaw cycles using a dry-ice/ethanol bath (-70 ^{0}C) and 37 ^{0}C incubation block. Vortex mixing in between these cycles aided cell lysis and disaggregated clumped cells. After the final cycle, the tube contents were centrifuged at 11,600 x g for 1 minute in order to pellet cell debris, and the supernatant collected. Total cell protein extracts made in this way were stored at -20 ^{0}C until use.

(b) Mouse and sheep tissue protein extracts : A crude total cellular protein extract was made from mouse and sheep tissue samples by manual homogenisation of milligram quantities of tissue in an eppendorf tube using a tapered, ground glass pestle and 3 volumes of Tris-PLP/PMSF buffer per milligram of tissue. Disruption of the tissue was completed by three freeze-thaw cycles (as described above). After a one minute centrifugation at 11,600 x g, the supernatant fraction was collected and the protein concentration determined (see below). A 5 mg/ml working stock was made and this extract stored at -20 ^{0}C until use in the assays. The remainder was stored at -80 ^{0}C.

(c) LT2 Salmonella extract : prepared as a whole cell protein extract by A. Sivaprasad. LT2 bacteria grown under the conditions described by Kredich and Tomkins (1966) were lysed by sonication, the cell debris removed by centrifugation for 1 minute at 11,600 x g, then the supernatant used as a crude positive control protein extract for the SAT and OAS sulphydrylase assays.

(d) GP001 Yeast extract : prepared by F. Lim according to the method of Young et al. (1969) in which cell lysis was achieved by disruption using the French press. Cell debris was removed by centrifugation as in (c) and the supernatant used as a crude positive control extract in the OAS sulphydrylase assays conducted in the MET25 work.

Estimation of protein concentration : The concentration of proteins in the crude extracts prepared (described above) were measured by the U.V. absorbance

method of Layne (1957) and the protein concentration calculated according to the formula...

Protein concentration. $= (A_{280nm} \times 1.55) - (A_{260nm} \times 0.76)$

(If the dilution factor is taken into account, the concentration given by this is in mg/ml.)

This method was found to be especially suitable because in most cases, minimal amounts of protein were available for analysis. Careful note of the nucleic acid content of such extracts was taken ($A_{280nm} : A_{260nm}$ ratio) as excessive amounts can contribute to error in the measurement of protein concentration using this method. As the method is known to be unreliable at concentrations at or below 0.1 mg/ml, extracts from small tissue samples were made in minimal volumes of the Tris-PLP/PMSF buffer.

Western transfer analysis. : Analysis of proteins from the MET25 RRL translation experiments (see section 3.4.1) and pLT-MET permanent CHO cell lines expressing the MET25 gene product (see section 3.4.2) fractionated in SDS-polyacrylamide gels by the Western transfer method was performed as described by Burnette (1981). Proteins in the gel were transferred to Nitrocellulose membrane (Sartorius) by electrophoresis at 400 mA (a 55 V potential difference was created between the electrodes, set 2.5 cm apart in the apparatus in use) for 3-5 hours. Following a 4 hour prehybridisation step (Western buffer 1 in materials) hybridisation of a MET25-specific antibody proceeded for 16 hours (a 1 in 5000 dilution (in buffer 1) of the antiserum prepared by Yamagata (1984) was used as the primary antibody). Five post-hybridisation washes (buffer 2) preceded a one hour hybridisation of the secondary antibody (a 1 in 1000 dilution (in buffer 2) of a goat anti-rabbit IgG/Alkaline phosphatase conjugate; No. A-8025, Sigma) to the membrane. Five further washes (buffer 2) were then followed by application of the chromogenic substrate Nitroblue tetrazolium / 5-bromo-4-chloro-3-indolyly phosphate (NBT/BCIP) colour detection system (in buffer 3) described by Forster et al. (1985). The colour reaction was allowed to proceed in darkness until satisfactory visualisation of the sites of binding of the primary antibody to proteins initially transferred to the membrane was achieved, then stopped by addition of buffer 4. A photographic record of the result was then taken.

Assays of enzyme activity in protein extracts.

(a) Serine acetyltransferase (SAT) assay : The SAT assay used during the course of the work described in this thesis was based on the method described by Kredich and Tomkins (1966). In their assay, ^{14}C-serine was used as the substrate and the acetylation of this to form OAS was carried out at 25 ^{0}C in a volume of 0.5 ml, including 1.0 mM L-serine, 0.1 mM acetyl-CoA, 1.0 mM EDTA, 50.0 mM Tris.HCl (pH 7.6) and an appropriate amount of enzyme. Following incubation, an aliquot was spotted onto a TLC sheet, air dried and chromatographic separation of reactants and products achieved using the solvent chloroform: ethanol: glacial acetic acid: water in the ratio 50 : 32 : 10 : 8. The dried chromatogram was then subjected to autoradiography at room temperature.

Many features of this assay method have been altered to better suit the simultaneous handling of multiple small samples; the reasons for this are as stated in the case of detection of OAS sulphydrylase activity (see section 3.3). The assay devised for use in the current work is performed in eppendorf tubes in a 20 μl volume at 37 ^{0}C (closer to that of mammalian cells in normal growth conditions, this temperature also facilitates greater catalytic activity but does not lower enzyme stability) and includes 30 μM ^{14}C-L-serine (a concentration sufficient to allow proper enzyme function and to enable detection of small quantities of active enzyme; supplemented with 1.0 mM L-serine when quantitation of enzyme activity is attempted) 1.0 mM acetyl-CoA, 5.0 mM Tris.HCl (pH 7.5) (sufficient to maintain the optimum pH of the reaction mixture but within the tolerance limits of the TLC procedure in regard to solute concentration; reactions have also been performed sucessfully at pH 6.5 - 7.0) and up to 20 μg of test protein from a crude extract (added last to begin the reaction). After a 30-60 minute incubation, each sample is heated at 100 ^{0}C for 2-3 minutes, denatured protein removed by a 1 minute centrifugation and 5 or 10 μl of the supernatant (one quarter to one half of the reaction mixture) spotted onto a TLC sheet. The chromatographic separation of reactants and products in the solvent described above (or in the Checchi solvent, used in the OAS sulphydrylase assay procedure) over a development distance of 10 cm is followed by drying of the TLC sheet and autoradiography for 1-3 days. Some alteration of this method has been necessary to increase the sensitivity of detection and where relevant, the details of such

changes will be given. Nonetheless, reliability, repeatability and high sensitivity are features of this method which have lent it to extensive use in this work.

To begin an analysis of transgene expression in vivo, the SAT assay method was employed to detect active enzyme in protein extracts made from the tail tissue, and later from other tissues, of the transgenic founder mice. The rapidity, sensitivity and convenience afforded by this method were the main reasons for its use. The method of preparation of protein extracts from mouse tissues is given above. Though similar to other reported methods, this does not include a 65 ^{0}C heat treatment step (usually included to facilitate uninhibited use of acetyl-CoA by a particular test enzyme, e.g. CAT; Sleigh, 1986) because SAT enzyme held at this temperature was shown to be heat labile and to lose greater than 50 % of its activity after 5 minutes (data not shown). In addition, the protease inhibitor phenylmethylsulphonyl fluoride (PMSF) was added to the tissue protein before homogenization, to a final concentration of 1 mM, to overcome the high level of SAT-degrading protease activity detected in many tissues (data not shown).

(b) OAS sulphydrylase (sulphydrylase) assay : A typical assay, performed in a 20 µl volume (eppendorf tube) at 37^{0}C, contained 10 µg of total cell protein (e.g. from a crude extract of the wild-type strain GP001) as a source of the sulphydrylase enzyme and included 12.5 mM ^{14}C-OAS, 25 µM pyridoxal phosphate, 10 mM Potassium phosphate (or Tris) buffer (pH 7.8) and 17.6 mM 4-vinylpyridine (a three-fold molar excess over the theoretical maximum of cysteine which could be produced). To begin the reaction, sodium sulphide was added to a final concentration of 5.0 mM.

Following incubation of the reaction mixture at 37^{0}C for 15-60 minutes, 2.5 volumes (50 µl) of room-temperature ethanol was added to precipitate any protein and nucleic acid present. After thorough vortex mixing, the precipitate was pelleted by a 2 minute centrifugation (Microcentaur, M.S.E.) at 11,600 x g. Half of the ethanol supernatant (35 µl; containing unreacted ^{14}C-OAS and 4-vinylpyridine, other substances soluble in 70% ethanol including the buffer ions and sodium along with any cysteine-4VP formed) was transferred into a new eppendorf tube and vacuum dried (Speed Vac Concentrator, Savant) for 30 minutes. The dried solutes were resuspended in 5 µl of H$_2$O by vigorous

vortexing and then loaded directly onto a TLC sheet (Silica gel,Merck). The sample spots were dried and the substances present fractionated by ascending thin layer chromatography in a sealed glass chamber using the solvent 1-butanol/acetic acid/water in the ratio 12:3:5 as described by Checchi (1961). Finally, the chromatogram was dried and the ^{14}C-labelled products detected by autoradiography at room temperature.

(c) Cysteine synthesis assay : The cysteine synthesis assay, used to detect the enzyme activities of cysE and cysM gene products found in the same protein sample, contained a combination of the components of the SAT and OAS sulphydrylase assays, namely 30 μM L-^{14}C-serine, 1.0 mM acetyl CoA, 5.0 mM Tris.HCl (pH 6.5), 17.6 mM 4-vinylpyridine and 5.0 mM sodium sulphide. In a typical assay in a 20 μl volume, 20 μg of protein from a crude extract was combined with a mastermix containing all of the above (prepared fresh, and with sodium sulphide added just prior to dispensing of mastermix aliquots) in an eppendorf tube, the lid firmly sealed (to prevent the escape of H_2S and 4-vinylpyridine vapours) and the assay carried out at 37 ^{0}C for 30-60 minutes. At the conclusion of the reaction, each sample was processed, fractionated by TLC and autoradiographed as described for the OAS sulphydrylase assay samples above. The result of a cysteine synthesis time course assay is shown in figure 5.7 and demonstrates the expected shift in radioactivity (^{14}C) from serine to OAS, then from OAS to cysteine.

(d) Preparation of ^{14}C-labelled OAS and NAS : The mono-acetylated forms of serine, namely OAS and NAS, used as reference compounds in the TLC runs and as a substrate (OAS only) in the cysteine synthesis assays (particularly in the MET25 work described in Chapter 3) were prepared by the method of Greenstein and Winitz (1967).

2.2.4 Cell culture methods.

Maintenance of cell cultures. : Where sterile conditions and technique were required to maintain viability, cell culture handling in the methods described below was routinely performed in a biohazard safety hood (BH series, Class II, Gelman Sciences Australia Pty. Ltd.) and according to the C1 containment handling guidelines specified by the Biohazards Committee (Univ. of Adelaide).

(a) Storage of cell lines : Cell line stocks (cells from a 50-70 % confluent monolayer in a T75 culture flask) prepared for use in the cell culture experiments described (i.e. CHO, COS-1, HeLa, RESV40 and CSL503) were harvested by trypsinisation (see below) then kept stored in liquid nitrogen after resuspension in 1.5 ml 10 % DMSO in foetal calf serum (FCS). Initial freezing to -80 ^{0}C was followed by transfer of the frozen stock tubes to the liquid nitrogen.

(b) Recovery of cell cultures from frozen storage : Cells stored frozen in 1.5 ml stocks (10 % DMSO in FCS) were thawed rapidly to 37 ^{0}C, combined with 8.5 ml of the appropriate culture medium containing 10 % FCS, pelleted gently by centrifugation (1,000 revolutions per.minute (r.p.m.)/ room temperature in a Hettich "UNIVERSAL" bench centrifuge) the supernatant discarded and the cells resuspended in 10 ml of the appropriate culture medium (containing 10 % FCS) and seeded into a culture flask (equivalent in size to that from which the cells were derived prior to storage). When in culture medium, cell cultures were kept in a 37 ^{0}C / 5 % CO_2 / humidified environment in a water jacketed incubator (Forma Scientific).

(c) Passaging of cells in culture : CHO cell transformants maintained in culture for extended periods under G418 selection were routinely passaged after flasks had reached confluence, by trypsinisation and reseeding of the flasks at a low cell density in Ham's F12 medium / 10 % FCS supplemented with 350 μg/ml G418. All other cells were maintained by culture in DMEM / 10 % FCS and incubated under the conditions described above. Gentamicin ("Garamycin", Schering Corp., USA) was included in all media at a final concentration of 50 μg/ml.

(d) Harvesting of cells for DNA / RNA / protein : Routine harvesting of cells in monolayers from flasks (or dishes) prior to extraction of DNA, RNA or protein was performed by first decanting the culture medium, washing monolayers twice in phosphate buffered saline (PBS) and detaching the cells by treatment with a trypsin-EDTA solution (enough of a 1.0 mg/ml trypsin in 1 x "Versene" to cover the cells in the monolayer) at room temperature for 2-3 minutes (or until cells became visibly rounded and were easily detached from the flask by gentle tapping on the underside of the flask). Culture medium containing 10 % FCS was then added to inactivate the trypsin, the detached cells aspirated

using a pipette until a homogeneous cell suspension was obtained and cells harvested by centrifugation at 1000 r.p.m. as described above. Further washes and resuspension of the cells were performed using the buffers appropriate to the intended purpose of the cell material. In the case of preparation of RNA from cultured cells, handling times during harvesting were minimised.

(e) Selection of G418-resistant CHO transformants : Following transfection of CHO cells with plasmids expressing the aminoglycoside phosphotransferase gene (see Chapter 3, section 3.4.2) by electroporation (see below) the cells were seeded into 60 mm dishes, maintained in Ham's F12 medium / 10 % FCS without G418 for 3 days, then the medium replaced with the same type supplemented with 350 µg/ml G418 ("Geneticin", Gibco Laboratories). Medium was replaced every 3 days from this point, until the "No DNA" control transfection plate contained no viable cells (usually after about 7 days). After 7-14 days in these culture conditions, colonies of cells derived from single CHO cell transformants (i.e. individual cell clones) became established in the test plates. These were transferred to separate 2 ml wells in multi-well trays, maintained under G418 selection and the separate cultures expanded until sufficient cells were available to create frozen stocks. Cultures were then expanded again for analysis at the DNA, RNA and protein levels. The cloning procedure, utilising sterile glass cloning rings to facilitate isolation was carried out essentially as described by Freshney (1987).

(f) Cell counting prior to transfection : Estimates of numbers of viable cells from harvested cultures (homogeneous cell suspensions) were performed using a haemocytometer (Neubauer, Improved) and a preparation of 0.5 % trypan blue in PBS. Cells were diluted 1 in 10 in this solution prior to cell counts.

(g) Preparation of cell culture media and reagents : Cell culture media used in this work were prepared by filter sterilisation (0.2 um, Sartorius) using sterile glassware and filtration apparatus in a laminar flow hood (Gelman Sciences) and were made with tissue culture grade water (final filtration with a 0.2 µm filter). All other reagents required for the cell culture experiments were prepared using sterile technique, as described by Freshney (1987).

Transfection protocols for transient expression studies.

(a) DEAE-Dextran transfection (HeLa and RESV40 cells) : Transient transfection of HeLa and RESV40 cells was performed according to the method of Lopata et al. (1984) using 500 µg/ml DEAE-Dextran in a 1 hour transfection step and 30 % DMSO in a 2 minute "shock."

(b) Electroporation (CHO, COS-1 and CSL503 cells) : Transfection of CHO, COS-1 and CSL503 cells for transient expression studies was achieved using the electroporation method described by Chu et al. (1987) and the "GENE PULSER" electroporation apparatus (BIORAD). A table detailing the conditions employed in the case of these cell lines is given below. Cells in culture were harvested by trypsinisation 48 hours after transfection using these protocols.

Cell type	No.of cells per experiment	Buffer conditions	Buffer volume	Pulsing capacitance/voltage (ufd / V)
CHO	5×10^6	PBS	0.8 ml	25 / 1300
COS-1	5×10^6	HepesBS/FCS/glucose/ss DNA	0.5 ml	250 / 300
RESV40	5×10^6	HepesBS/FCS/ glucose/ssDNA	0.5 ml	960 / 380
CSL503	5×10^6	HepesBS/FCS/glucose.	0.5 ml	960 / 270

* 2-10 ug of DNA was used per transfection, and all cell lines used were adherent lines.
** ssDNA is salmon sperm DNA (1.0 mg/ml final), HeBS is Hepes buffered saline.

Transfection of CHO cells to produce permanent cell lines. : To create stably transformed G418-resistant CHO cell lines, transfection of appropriate plasmid constructs was again carried out by electroporation, but using only 5×10^5 cells per transfection. All other transfection conditions were as described for electroporation of CHO cells in transient expression experiments.

2.2.5 Transgenesis Methods.

2.2.5.1 Production of transgenic mice.

The following methods were performed by Mr. Paul Verma (and co-workers in the Department of Obstetrics and Gynaecology, University of

Adelaide, South Australia) essentially according to the methods described by
Brinster et al. (1981) and Hogan et al. (1986).

(a) Maintenance of the C57BL/6J mouse strain.

(b) Preparation of pseudopregnant females recipients.

(c) Preparation of donor females / vasectomised males.

(d) Collection and fertilisation of ova.

(e) Microinjection and transfer of embryos.

Preparation of DNA fragments for microinjection.

(a) The cysE transgene : The 3.1 kb SalI / ScaI cysE transgene fragment
was excised from pLT-E and prepared for microinjection by preparative agarose
gel electrophoresis, transfer to and elution from NA-45 ion-exchange paper
(Schleicher and Schuell) and ethanol precipitation. The DNA was lyophilised,
resuspended in PBS at 20 ng/ul, then diluted further in PBS and centrifuged at
11,600 x g for 2 minutes prior to pronuclear microinjection. Approximately 3000-
6000 copies of the fragment were delivered to the male pronucleus in 2 pl during
microinjection of fertilised ova.

(b) The cysE/cysM linked-transgene : The 6.4 kb SalI / ScaI linked-
transgene fragment was excised from pLT-EM and prepared for microinjection by
preparative agarose gel electrophoresis and purification via elution from glass
powder (Gene-clean, BIO 101). DNA prepared by this method was then
lyophilised and treated as described for the cysE transgene fragment prior to
pronuclear microinjection. This DNA was also used in pronuclear microinjection
of sheep ova.

Identification of founder transgenics.

(a) Enzyme assays : SAT, OAS sulphydrylase or cysteine synthesis
assays of tail protein extracts were used in initial identification of founder
transgenic mice. Whilst a rapid, convenient method of analysis, only transgenic
mice expressing the transgene could be identified.

(b) DNA analysis : DNA dot blots and Southern transfer analyses were
used to positively identify founder transgenic mice.

Analysis of gene expression in tissues of transgenic mice. : The analysis of gene
expression in tissues of transgenic mice was achieved by use of the SAT, OAS

sulphydrylase and cysteine synthesis assays of protein extracts made from selected tissues (described in section 2.3).

2.2.5.2 Production of transgenic sheep

All of the methods utilised in the production of the transgenic sheep described in this thesis were carried out as detailed by Walker et al. (1990). Sheep used in this programme were of the South Australian Merino breed. Identification of founder transgenics. : The enzyme assay and DNA dot blot and Southern transfer methods applied to the mouse transgenesis were also used to identify transgenic sheep, except in the case of an aborted lamb, where the PCR method was used to analyse the highly degraded DNA.

Analysis of gene expression in tissues of transgenic sheep. : The SAT and cysteine synthesis assays and the RNA protection analysis method were used to detect transgene expression in the tissues of transgenic sheep.

Containment and maintenance of transgenic animals. : The techniques of production, housing and maintenance of the animals produced in the course of the mouse and sheep transgenesis programmes described in this thesis conformed to the guidelines specified by the Genetic Manipulation Advisory Council and administered by the Animal Ethics and Biohazards Committees (Univ. of Adelaide) and were carried out in accordance with the provisions of the AEC permits M2/87A and M/62/88 (mouse transgenesis) and M/63/88 (sheep transgenesis) granted to Dr. R.F. Seamark.

2.2.6 Computer-assisted analysis methods.

Analysis of codon usage (see section 3.5.1) : performed with the aid of the "ANALYSEQ" programme devised and described by Staden (1984).

Analysis of secondary structure in RNA (see section 3.5.1) : carried out utilising the "Fold" programme of Zuker and Stiegler (1981) contained in version 5.0 of the Sequence Analysis Software Package of the University of Wisconsin Genetics Computing Group (GCG) described by Devereux et al. (1984).

* The input data analysed by these two computer programmes was processed by the Vax 11-785 computer.

Quantitative analysis of autoradiograms by densitometry : performed using the Molecular Dynamics Computing Densitometer (Model 300A) and "ImageQuant"

(version 2.0) analysis programme. (N.B. Pre-flashing of the autoradiographic emulsion was routinely carried out prior to autoradiography which was intended for quantitation. To do this, X-ray film was held 0.5 m from an Ilford (LR 915 filter) safelight for 15-30 seconds prior to exposure to radioactive samples.)

Chapter 3

Expression of the MET25 Gene.

Chapter 3 : Expression of the MET25 Gene.

3.1 Introduction

The yeast gene, MET25, was chosen as a candidate transgene; one which would provide an enzyme needed to catalyse the second reaction step in the pathway of cysteine synthesis, that being the conversion of O-acetyl-L-serine (OAS) to L-cysteine. The physicochemical properties of the enzyme encoded by the MET25 gene have been extensively investigated (Yamagata, 1987). Briefly, it is a 200 kD tetramer comprised of identical subunits, each of which contains one sulphydryl group and binds one molecule of an essential cofactor, pyridoxal phosphate. As mentioned in 1.6.1, it is a bifunctional enzyme known as O-acetyl-L-serine-O-acetyl-L-homoserine (OAS-OAH) sulphydrylase and when purified to near homogeneity, both of these enzyme activities may be detected. The ability of this enzyme to function independently of other proteins was a key factor in its choice for use in this work.

The MET25 gene encodes a single, 50 kD subunit of the tetramer. The isolation and initial characterisation of the gene was carried out independently in this and another laboratory, and is described elsewhere (D'Andrea et al., 1987). Its identity was confirmed by complementation analysis using strains mutant in this step of the yeast methionine biosynthesis pathway. The main body of work presented here describes the experimental determination of the suitability of this gene as a transgene for use in the sheep, largely by way of optimisation of expression of the gene in mammalian systems.

Considering the task at hand, many parameters pertaining to expression of the MET25 gene, a yeast gene, in a mammalian environment, had to be examined. These included (i) efficiency of translation of MET25 gene transcripts into protein, (ii) the activity of the MET25 gene product when it is synthesized outside the yeast environment and (iii) expression of the gene via transcription from a heterologous promoter (selection of a suitable promoter of transcription has already been discussed in relation to the need for tissue-specific expression). As far as possible, these facets of gene expression will be discussed separately.

Results.

3.2   Efficiency of translation of MET25 gene trancripts.

In order to assess the efficiency of translation of MET25 transcripts, two in vitro methods were employed in combination. MET25 transcripts were produced using the SP6 RNA transcription system, and protein was produced using the rabbit reticulocyte lysate (RRL) translation system. The use of the RRL provided a suitable mammalian environment in which translation of the yeast protein could be monitored. Translations were performed with RNA concentrations at or below the maximum reported for purified RNA species using this system (20.0 µg/ml; Pelham and Jackson, 1976; Van Tol and Van Vloten-Doting, 1979) and usually over a 10-fold range of concentration (2.0 - 20.0 µg/ml) so as to eliminate concentration-dependent effects and allow more accurate quantitation of translation efficiency. Comparisons could ultimately be made on the basis of assessing the performance of equal quantities of various transcripts in vitro in the RRL. Initially, the base vector used to provide the SP6 promoter sequences was the cloning vector pIBI76. The cloning of the gene into this vector is described in figure 3.1. The methods of synthesis of SP6 RNA transcripts and of use of the RRL to produce protein (Krieg and Melton, 1984; Pelham and Jackson, 1976; Clemens, 1983) are described in Chapter 2.

The rationale of initial experiments was that the MET25 SP6 transcripts be translated in the RRL translation system alongside SP6 transcripts of sequences known to be translated well in this mammalian system. This enabled a comparison of the efficiency of translation of the various sequences to be made. The other RNA species included the SP6 trancripts of the cysE gene of S. typhimurium and the CYS1 gene of S. cerevisiae. Both genes have been described previously (D'Andrea et al., 1989); the clones from which the SP6 transcripts for these genes are derived, are shown in figure 3.2. (The protein produced from translation of SP6 cysE gene transcripts in the RRL system was shown to have serine acetyltransferase activity.) Given that each of these transcripts had been shown to be translated well in the RRL system a comparison of translation efficiency with that of the MET25 SP6 transcripts was useful in providing an initial assessment of the MET25 translation efficiency.

Figure 3.1 : Construction of transcription vector pIBI-MET.

A 1485 bp SacI / ScaI fragment containing the entire MET25 gene and including 104 bp of 5'-untranslated and 52 bp of 3'-untranslated yeast sequence was isolated from the vector pSVMT-MET/A+ (i) (see Bawden et al., 1987) and subcloned directionally into SacI / SmaI digested pIBI76 (ii). The 5' termini created during the digestion of pIBI76 were dephosphorylated with Calf Intestinal Phosphatase (CIP; Boehringer) prior to ligation using T4 DNA Ligase (Bresatec). The vector created was pIBI-MET (iii). When this was linearised with BamHI, transcripts arising from the SP6 promoter were 1505 bases in length and provided the desired sense RNA of the MET25 gene. The message in SP6 RNA transcripts produced was then translated into protein in the Rabbit Reticulocyte Lysate Translation system. With reference to translation, the positions of the ATG initiation and TGA termination codons are shown.

(i)

SacI — ATG

BamHI

hMTIIa

EcoRI

HindIII

EcoRI

SV40 ori

pSVMT-MET/A+

6.68 Kb

MET25

HindIII

TGA

ScaI

PstI

amp

pBR ori

A+

BamHI

SalI

(ii)

EcoRI

SacI   SmaI

BamHI

SP6

lacZ

pBR ori

F1 ori

pIBI76

4.07 Kb

amp

ScaI

(iii)

EcoRI

SacI

ATG

EcoRI

lacZ

SP6

MET25

F1 ori

pIBI-MET

5.56 Kb

HindIII

TGA

ScaI/SmaI

BamHI

pBR ori

amp

ScaI
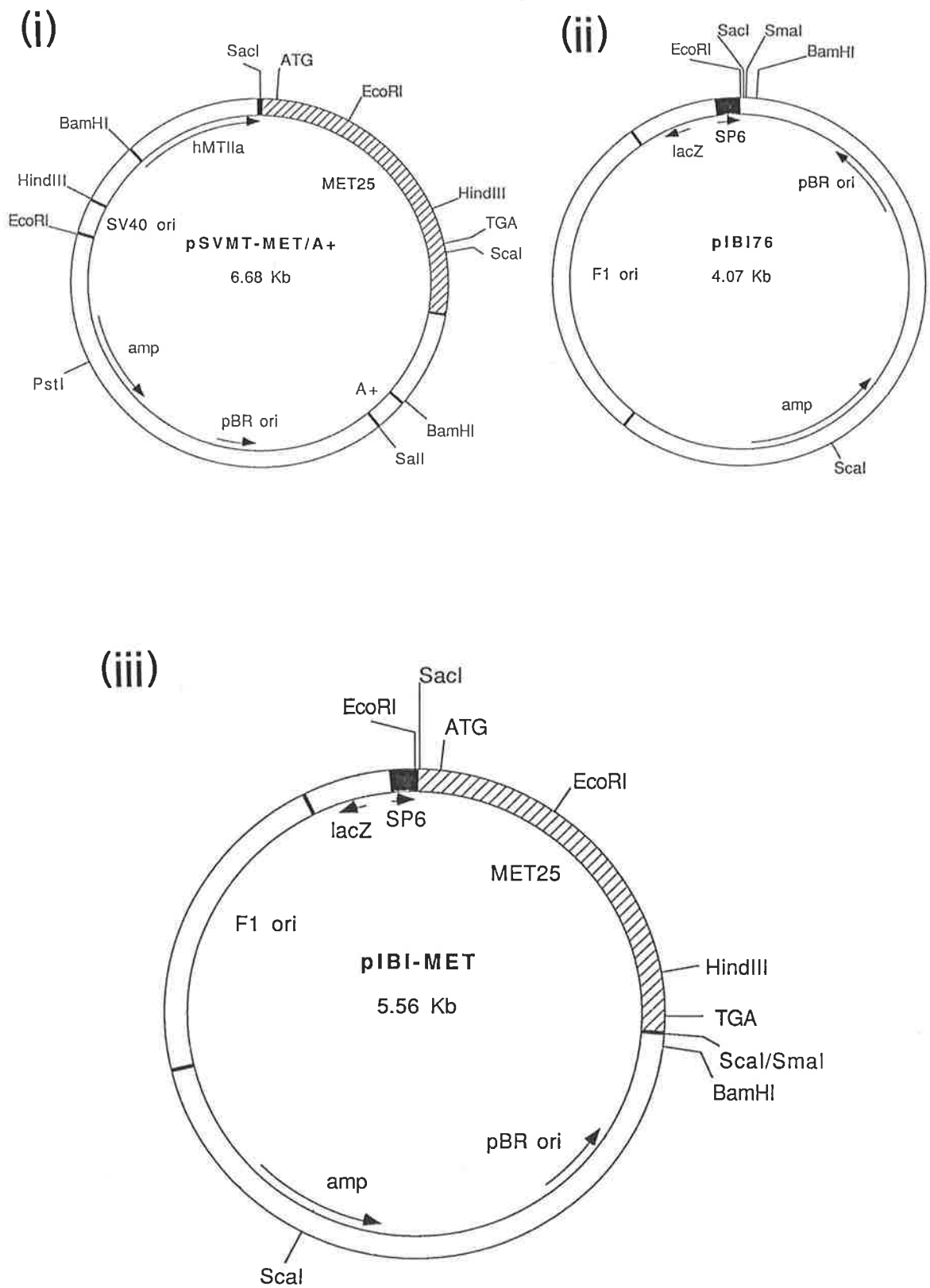
Figure 3.2 (a) : Construction of transcription vector pGEM-cysE.

A 979 bp DdeI / EcoRI fragment containing the entire cysE gene and including 89 bp of 5' non-coding and 68 bp of 3' non-coding sequence was isolated from the clone pRSM8 (Monroe and Kredich, 1988) and subcloned by blunt-end ligation into the SmaI site of the transcription vector pGEM-2 (Promega). The 979 bp insert was treated with the Klenow fragment of E. coli. DNA Polymerase I (KF) to blunt the termini prior to ligation. The SP6 promoted transcripts synthesized would be 1007 bases in length and provide a sense RNA from which a 34,000 M_r monomer of the cysE protein, serine acetyltransferase, could be translated in vitro. The resulting clone was pGEM-cysE (E. Kuzcek, unpublished) and was linearised with BamHI prior to in vitro transcription reactions. The positions of the ATG translation initiation and TGA termination codons are shown.

(b) : Construction of transcription vector pGEM-CYS1.

The CYS1 gene was directionally subcloned from the original vector pCYS101 into SmaI / HindIII digested pGEM-2, as a 2127 bp (Bal31-KF) blunt / HindIII fragment such that the coding region was flanked by 97 bp at the 5' end and 845 bp at the 3' end. Transcripts arising from the SP6 promoter were 1337 bases in length and used to provide sense RNA from which a 40,000 M_r monomer of the putative CYS1 protein, γ- Cystathionine Synthetase, could be translated in vitro. This clone was called pGEM-CYS1 (D'Andrea et al., 1989) and was restricted with AvaII prior to the SP6 transcriptions. Again, the positions of the ATG and TAA codons are shown.

a.

pGEM-cysE
3.83 Kb

EcoRI
ATG
SP6
cysE
TGA
BamHI
SalI
T7
amp
ScaI
pBR ori

b.

pGEM-CYS1
4.96 Kb

EcoRI
ATG
SP6
CYS1
TAA
AvaII
ScaI
amp
pBR ori
T7
HindIII
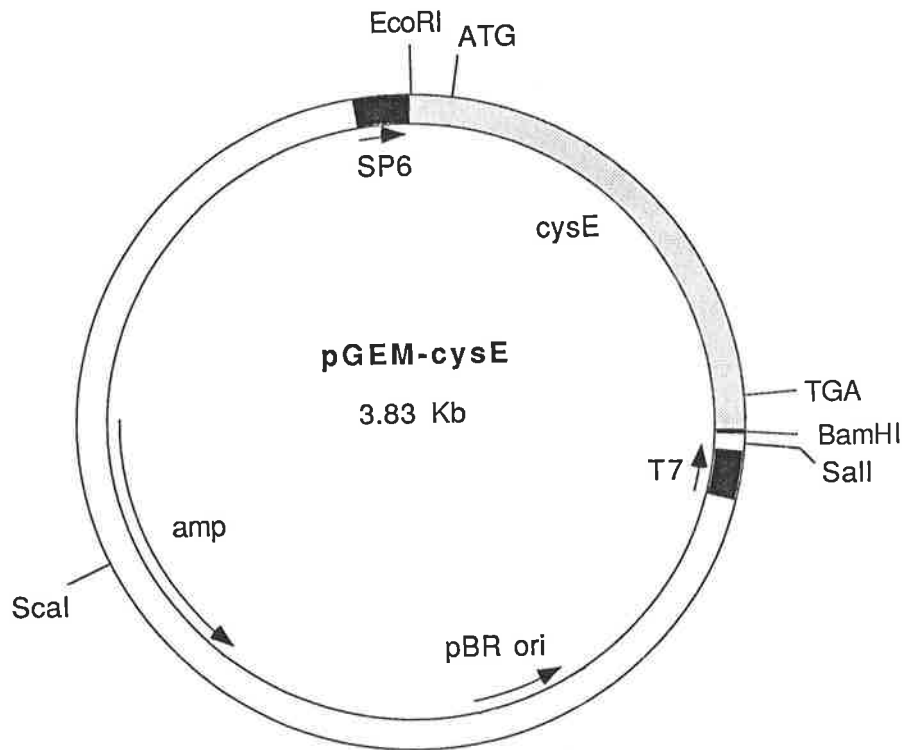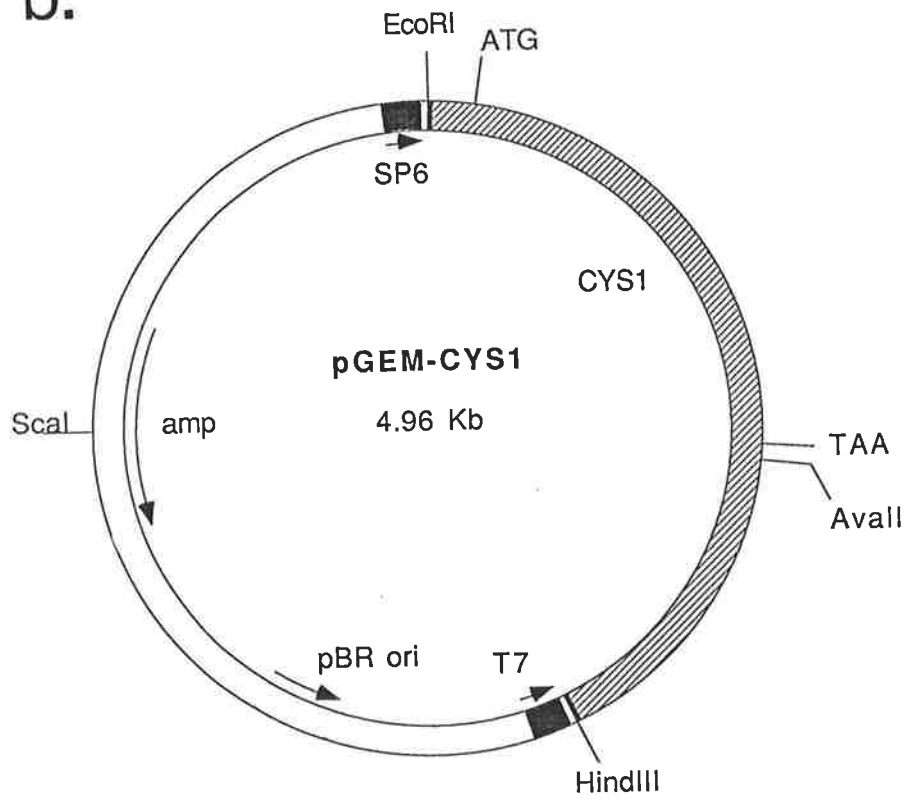
3.2.1 Initial characterisation of translation efficiency.

In these experiments, approximately equal amounts of uncapped SP6 transcripts were provided to the RRL system, and the translations performed in the presence of ^{35}S-methionine, together with an appropriate negative control containing no RNA. Radioactively labelled protein products were then fractionated using a denaturing SDS-polyacrylamide gel system and visualised by fluorography. These methods are described in Chapter 2. In repeated experiments, translation of MET25 SP6 transcripts was very poor when compared to the controls, cysE and CYS1, even taking into account the different proportion of ^{35}S-methionine which could be incorporated on the basis of the amino acid compositions of these protein species (figure 3.3). These results indicated that in contrast to cysE and CYS1, the efficiency of translation of MET25 gene transcripts in a mammalian environment was low and may be vastly improved.

It has been reported that translation efficiency is impaired by secondary structure in the 5'-untranslated region of the mRNA (Sleat et al., 1988). Interestingly, examination of the 5'-untranslated region of pIBI-MET using the "Fold" programme of Zuker (capable of predicting secondary structures that are likely to form in an RNA molecule; see Zuker and Stiegler, 1981; Frier et al., 1986) revealed that a large stem-loop complex involving 83 bases of this region (including bases from -21 to -104 with respect to the first base of the translation initiation codon at +1) was possible. It seemed probable that in order to optimise translation, removal of this segment of the yeast 5'-untranslated sequence would be necessary. Further, it has been suggested that there is an optimal sequence context for the initiating AUG codon (Kozak, 1987). These parameters (discussed again in section 3.5) were taken into consideration when a new series of gene constructs was created for use in the in vitro translation experiments. The constructs were designed with two specific aims in mind; (a) removal of extraneous yeast sequences (particularly in the 5'-untranslated region) which may be contributing to reduced translation efficiency, and (b) addition of heterologous sequences which have previously been demonstrated to improve translation efficiency in a mammalian environment and to otherwise stabilize transcripts

Figure 3.3 : Comparison of the translation of the MET25, CYS1 and cysE coding regions in the RRL.

In vitro translation of equal amounts (150 ng) of uncapped SP6 transcripts derived from the clones pIBI-MET (lane 2) pGEM-CYS1 (lane 3) and pGEM-cysE (lane 4) alongside a negative control containing no added RNA (lane 1) was performed using the RRL translation system (Amersham). The proteins were synthesized in the presence of L-[^{35}S] methionine and fractionated in a 12.5 % SDS-PAGE system. The distance of migration of molecular weight standards (BRL prestained markers; size given in kD) is indicated on the lefthand side of the figure, and the position of the major product synthesized in each RRL reaction is marked (►) on the left side of each lane. Fluorography of the gel was carried out at -80 ^{0}C for 16 hours.
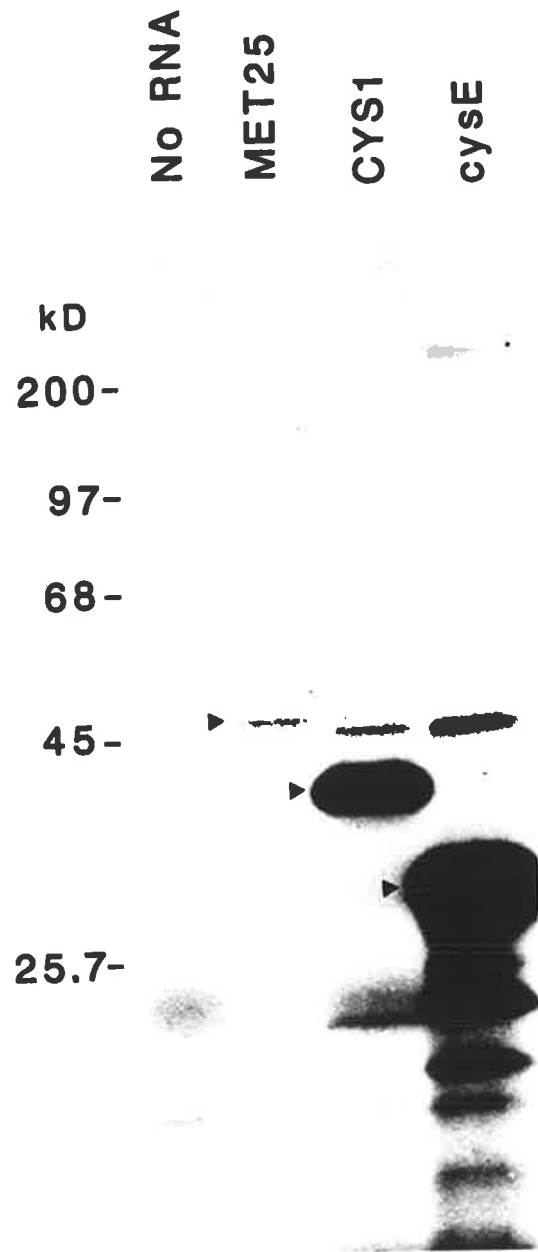
intended for use in such in vitro translation systems. In addition to (a) and (b), the requirement of a 5'm7G cap structure for efficient translation of the MET25 mRNA was tested and is described in (c) to follow.

3.2.2   Manipulation of the MET25 gene to improve translation.

(a) Removal of extraneous 5'-untranslated sequences : Whilst the nucleotide sequence of the coding region itself appeared to present no difficulty to the translation machinery (i.e. although some rarely used codons do exist within the first part of the coding region, a protein product of the correct size was consistently translated from SP6 transcripts of the MET25 gene in the RRL translation system) the 100 bp region immediately 5' to the initiation codon was thought likely to be contributing to reduced translation efficiency in the RRL system, and so was removed from pIBI-MET by deletion using Bal31 exonuclease to create pIBI-METΔ-4 (figure 3.4 a). This deletion left only 4 bp 5' to the initiation codon, and the sequence as T ACA ATG C. (This form of the MET25 gene, still containing 52 base pairs of sequence 3' to the termination codon, was used in all subsequent cloning events involving this gene.)

(b) Addition of the Xenopus ß-globin 5'-untranslated region : A sequence known to improve the translation of heterologous coding regions in the wheat germ extract and Xenopus oocyte translation systems is the 43 bp 5'-untranslated region of the Xenopus ß-globin gene. [Other 5'-untranslated sequences which enhance translation have been reported in the literature (Gallie et al., 1987a and b; Jobling and Gehrke, 1987; Sleat et al., 1988) and it is proposed that the greater translation efficiency of the mRNA species involved is due to the lack of secondary structure in these sequences (Jobling and Gehrke, 1987; Sleat et al., 1988).] When the 43 bp 5'-untranslated region is placed 5' to a heterologous coding region, SP6 transcripts produced are translated in these systems more efficiently than identical transcripts which lack this sequence (e.g. as is seen in the case of the   Xenopus ß-globin / human B-Interferon clone;   P. Krieg, pers. comm.), A similar improvement in translation efficiency by inclusion of this sequence in synthetic mRNA's might be expected in the RRL translation system, or when the mRNA from such a hybrid clone is processed in vivo. Hence, the cloning vector containing this sequence, pSP64-T (Krieg and Melton, 1984) was

Figure 3.4 (a) : Construction of transcription vector pIBI-METΔ-4.

Following SacI endonuclease digestion of pIBI-MET DNA, yeast sequences 5' to the ATG initiation codon of the MET25 gene coding region were removed using Bal31 exonuclease digestion; the resulting molecules were then treated with the Klenow fragment in order to blunt the termini. Deleted forms of the yeast sequences were excised with BamHI and fragments smaller than 1485 bp subcloned directionally into HincII / BamHI digested and CIP treated M13mp18. Deletion clones were compared by performing dideoxy "T" reactions on single-stranded DNA and the identity of deletion clone M13mp18 BalΔ-4, thought to represent a deletion of all but 4 nucleotides 5' to the MET25 gene initiation codon was confirmed by complete DNA sequence analysis. The sequence above the plasmid map shows the ATG underlined and the T residue at the -4 deletion point is marked (●). The PstI site and HincII half-site (GTC) from the M13 vector are also shown.

To construct pIBI-METΔ-4, the deleted form of the yeast MET25 gene, devoid of all except 4 bp 5' to the ATG codon, was excised from M13mp18 BalΔ-4 as a 1385 bp PstI / KF / BamHI treated fragment then subcloned directionally into SacI / KF / BamHI / CIP treated pIBI76.

a.

5'-AAGCTTGCATGCCTGCAGGTCTACAATGCCATCT-3'

PstI

SacI/PstI
EcoRI   ATG
                    EcoRI

lacZ   SP6

                    MET25

pIBI-METΔ-4
5.46 Kb

F1 ori   HindIII

                            TGA
                            BamHI

            pBR ori

    amp

ScaI
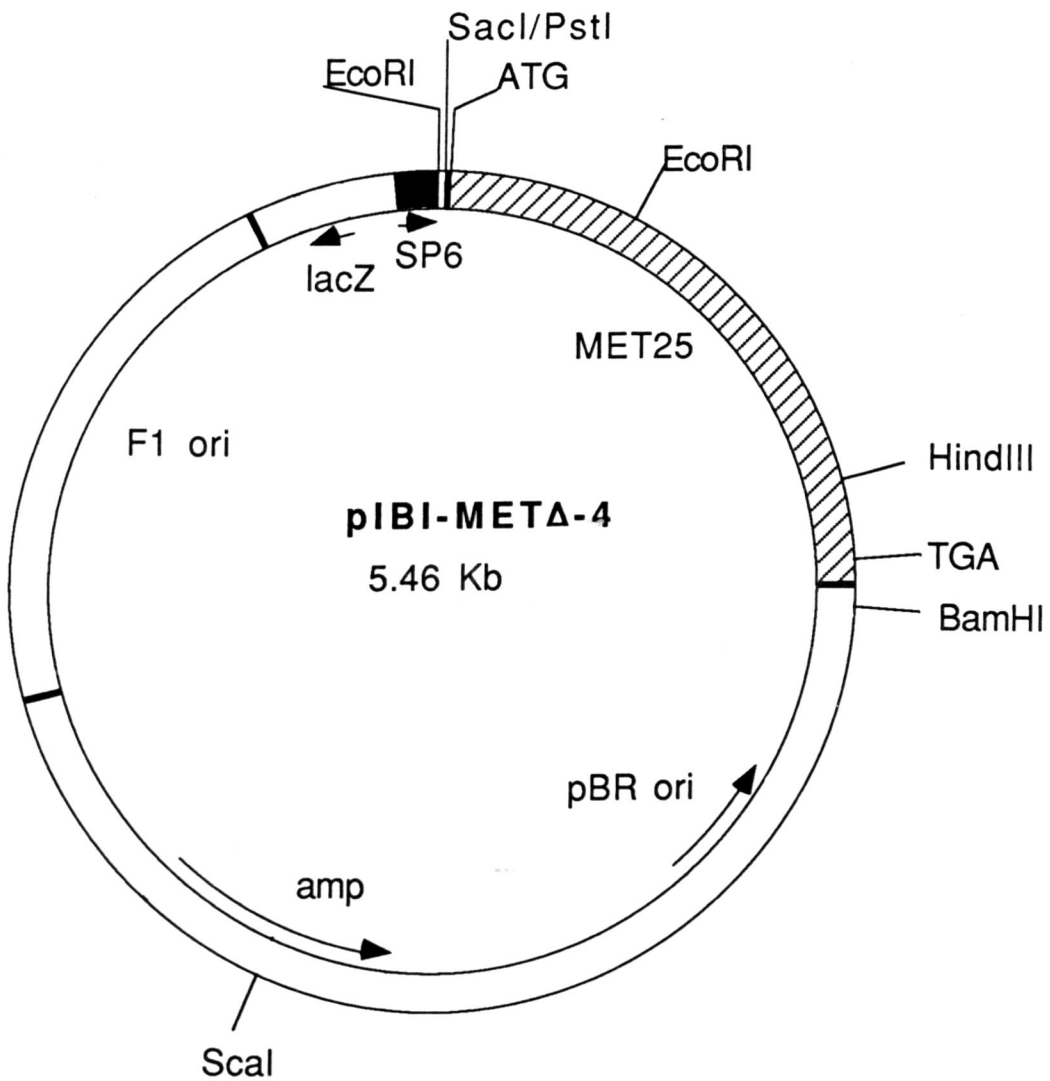
used as a base vector into which the full length and deleted forms of the MET25 gene (described above) were cloned. The structure of this base vector is such that coding sequences inserted 3' to the SP6 promoter and ß-globin 5'-untranslated sequences are flanked at the 3' end by 145 bases of <u>Xenopus</u> ß-globin 3'-untranslated sequence and a poly dA/dC tract of some 53 nucleotide residues, including 23 dA and 30 dC nucleotides. Whilst the 5'-untranslated globin sequences may contribute to increased initiation of translation, these 3' sequences are thought to confer a degree of stability on the SP6 transcripts produced. The new constructs made in this way were named pSPT-MET and pSPT-METΔ-4 (figure 3.4 b).

(c) Testing the requirement for a 5'm7G cap structure : Prior to a more rigorous examination and quantitation of the effects on translation efficiency of various trancripts, the effect of "capping" of the synthetic transcripts for use in such <u>in vitro</u> experiments was tested. Previous work has illustrated that the results of <u>in vitro</u> experiments more accurately reflect the <u>in vivo</u> situation if the 5'end of transcripts include a m7G(5')ppp(5')N cap structure, a common modification to the 5' termini of mRNA molecules in eukaryotes which is added to nascent RNA Polymerase II transcripts in the nucleus. The role of the cap structure has been documented to include facilitation of translation initiation, protection of the 5' end from exonucleolytic degradation and as being a structure recognized by ribonucleoprotein splicing complexes of transcripts containing introns (Konarska <u>et al.</u>, 1984).

To test the cap requirement directly in relation to the transcripts in use here, with a view to measuring translation efficiency under optimal conditions, SP6 transcripts from pSPT-METΔ-4, (trace-labelled with ^{32}P-rUTP for quantitation) were synthesized in the absence or presence of the cap analogue m7G(5')ppp(5')G (Biolabs) using a method described previously (Krieg and Melton, 1987). Transcripts were then translated in the RRL system in the presence of ^{3}H-serine (a 1.0 - 40.0 µg/ml range of concentration of RNA was used) and amounts of protein synthesized were quantitated by densitometry following fluorography. The result confirmed the requirement of a 5' cap structure for improved translation of MET25 transcripts in the RRL; when SP6 transcripts of pSPT-METΔ-4 contain a 5' cap structure, the efficiency of

Figure 3.4 (b) : Construction of transcription vectors pSPT-MET and pSPT-METΔ-4.

A 1485 bp insert was excised from pIBI-MET (i) by SacI / BamHI endonuclease digestion and the 5' termini blunt-ended by treatment with the Klenow enzyme; ligation of this to BglII digested and Klenow / CIP treated pSP64-T DNA (ii) created the vector pSPT-MET (iii) which still contained the 104 bp of 5'-untranslated and 52 bp of 3'-untranslated MET25 gene sequence. To create the vector pSPT-METΔ-4 (iv) a 1385 bp fragment was prepared from the deletion clone M13mp18 Bal^-4 by PstI / BamHI / KF digestion and subcloned into BglII / KF / CIP treated pSP64-T by blunt-end ligation. The new vector contained the deleted form of the MET25 gene, with only 4 bp of original yeast sequence immediately 5' to the initiation codon but still including 52 bp 3' to the TGA stop codon. The Xenopus β-globin 5'-untranslated and 3'-untranslated regions (shown as X. B-g. 5'-UT and 3'-UT) and associated sequences surrounding the BglII cloning site (see section 3.2.2.2) are shown in greater detail in the line diagram below. The position of the SP6 promoter 5' to these sequences and the poly dA/dC tract at the 3' end are shown. The last eight codons of the Xenopus β-globin gene (including the TAA codon) are contained in the 3' sequences and found only 5 bp 5' to the HinfI site.
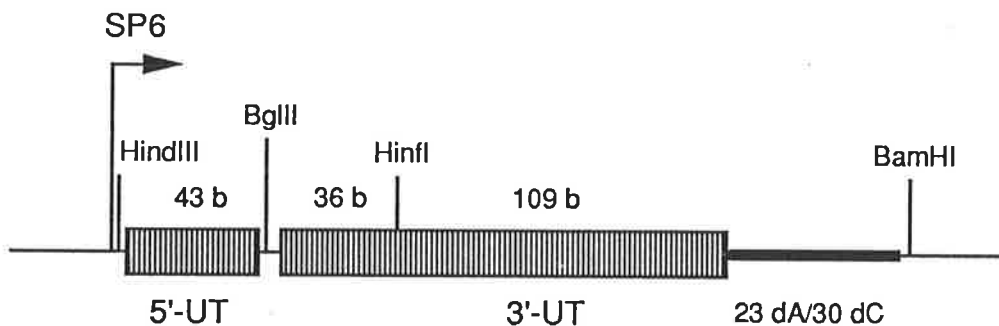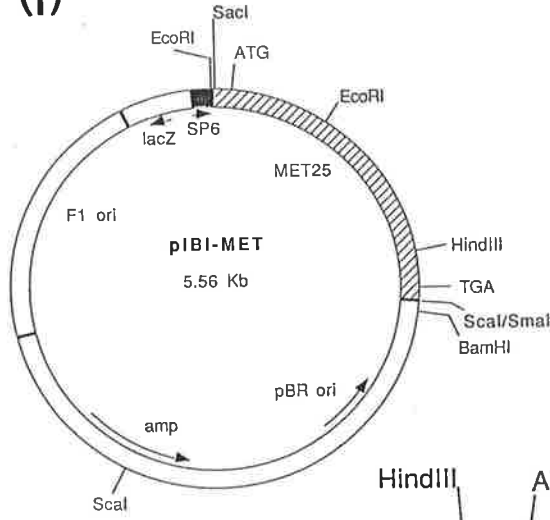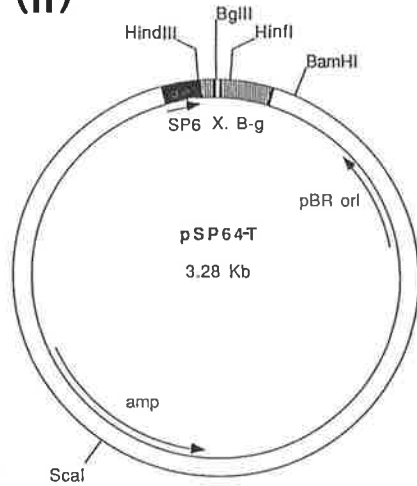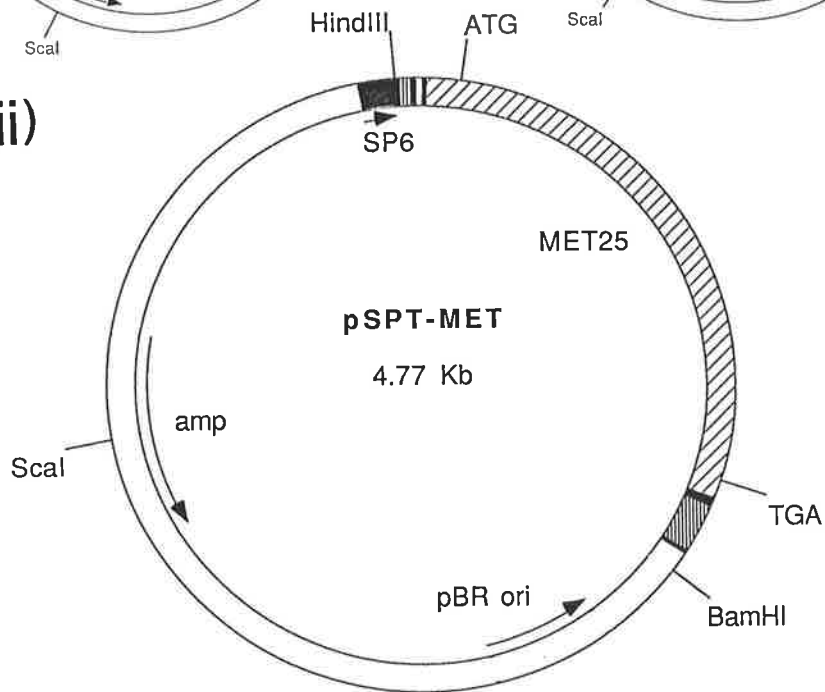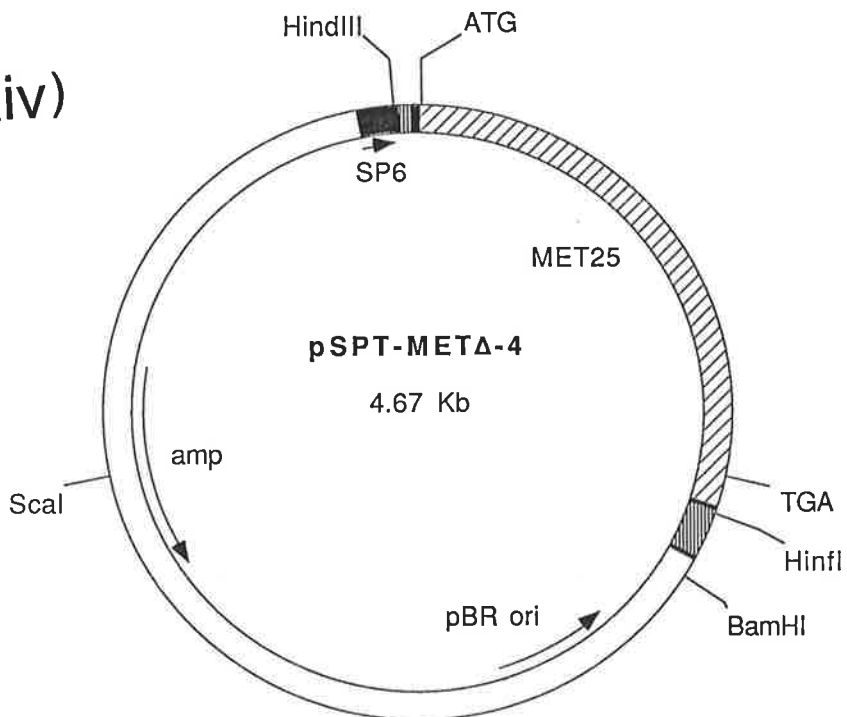
translation at 10.0 µg/ml of RNA is 0.75-fold higher than that of identical transcripts which lack this structure (data not shown).

Immediately following this, a direct comparison of the translation efficiencies of equal concentrations of capped SP6 transcripts derived from the clones pIBI-MET, pIBI-METΔ-4, pSPT-MET and pSPT-METΔ-4 was undertaken. In each case, translation of transcripts in the RRL was performed over the RNA concentration range of 2.0 - 20.0 µg/ml, with particular sample RNA concentrations being 2.0, 5.0, 10.0, 15.0 and 20.0 µg/ml. The amount of MET25 protein synthesized in these translations was compared by densitometry following visualisation of the tritiated protein products via fluorography. The results of this experiment are shown in figure 3.5 a and b, and the quantitative comparison of translation of the clones given in the table in figure 3.5 d. In brief, when translated in the RRL system, the ratio of the relative translation efficiencies of capped SP6 trancripts derived from the clones pIBI-MET, pIBI-METΔ-4, pSPT-MET and pSPT-METΔ-4 respectively is 1.0 : 0.89 : 1.94 : 2.48 when expressed as a comparison of the optimum level of translation obtained at an RNA concentration of 10.0 µg/ml.

3.2.3 Effect of removal of the 5'-untranslated region of MET25.

Despite removal of the 100 bp 5'-untranslated region from the MET25 gene in the clone pIBI-MET, the translation efficiencies of transcripts derived from this and pIBI-METΔ-4 are similar, with that of transcripts derived from pIBI-METΔ-4 actually being a little lower than those derived from pIBI-MET. In contrast, there is a significant difference between the translation efficiencies of transcripts derived from the corresponding pSPT-MET and pSPT-METΔ-4 clones. In this case, transcripts derived from pSPT-MET, which still contains the extra 100 bp of yeast sequence, are translated less efficiently than those from the pSPT-METΔ-4 clone.

3.2.4 Effect of addition of the Xenopus ß-globin 5'-untranslated region (and associated 3'-untranslated sequences)

When the level of translation of transcripts derived from either of the pIBI-based clones is compared with the same for either of the pSPT-based

Figure 3.5 : Comparison of the translation efficiencies of MET25 transcripts derived from (a) pIBI-MET and pIBI-METΔ-4 (b) pSPT-MET and pSPT-METΔ-4, and with the transcripts derived from (c) pGEM-cysE and pGEM-CYS1.

Following translation of each of the respective transcripts in the RRL at RNA concentrations of 0, 2.0, 5.0, 10.0, 15.0 and 20.0 μg/ml in the presence of L-[^{3}H] serine, the tritiated products were fractionated by SDS-PAGE and the gels fluorographed together at -80 ^{0}C for 60 hours. Translation samples containing no RNA are on the left side of each gel (marked -) and the remaining samples (with the RNA concentrations listed above) were loaded from left to right. The position of the major protein product in each sample is indicated (◄). The resulting autoradiograms were subjected to densitometric analysis and a quantitative comparison of the translation efficiencies of the transcripts is shown in the table in (d). In the table, the presence of the MET25 or Xenopus β-globin 5'-untranslated region is indicated along with the total length of the 5'-untranslated region of each clone. In addition, the calculated efficiencies take into account the serine composition of the different protein species (i.e serine in the MET25 protein product versus that in the cysE and CYS1 protein products).

a.   —   pIBI-MET   pIBI-METΔ-4

b.   —   pSPT-MET   pSPT-METΔ-4

c.   —   pGEM-cysE   pGEM-CYS1
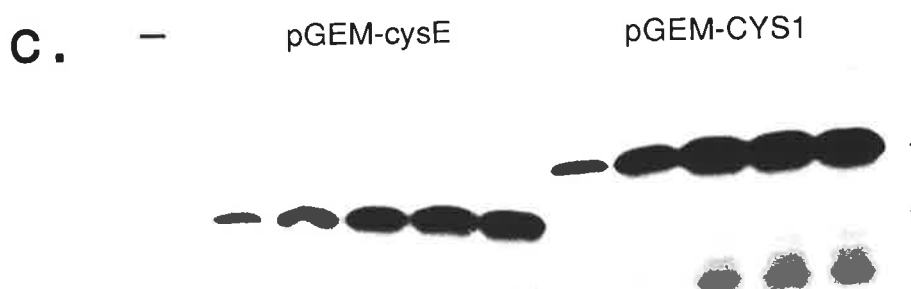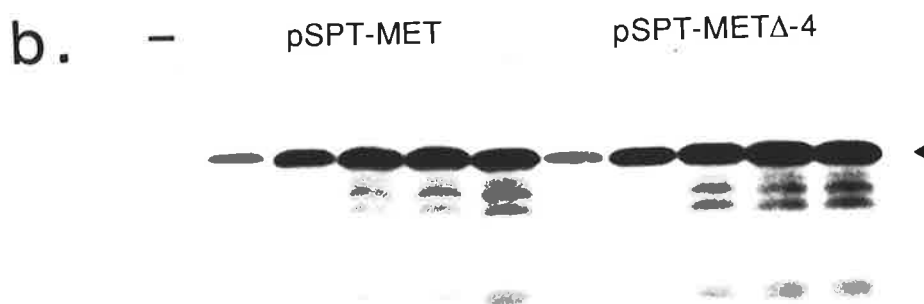
d.

Transcription clone	MET25 5'-UT region (100 bp)	Xenopus B-globin 5'-UT region (43 bp)	Total length of the 5'-UT region	Translation efficiency in the RRL
pIBI-MET	+	-	118	1.0
pIBI-METΔ-4	-	-	19	0.89
pSPT-MET	+	+	169	1.94
pSPT-METΔ-4	-	+	74	2.48
pGEM-cysE	- (89 bp cysE)	-	114	2.69
pGEM-CYS1	- (81 bp CYS1)	-	106	4.02

clones, it is clear that more MET25 protein product is synthesized in the case of the latter clones, where the resulting transcripts contain the Xenopus ß-globin 5'-untranslated region and associated 3'-untranslated sequences. This is the case irrespective of the presence of the 100 bp of extraneous yeast sequence.

3.2.5 Comparison with transcripts known to be efficiently translated.

To place this work in the context of proteins known to be translated well in the RRL, the same experiments were carried out with equal quantities of capped SP6 transcripts derived from the clones pGEM-cysE and pGEM-CYS1. These results are shown in figure 3.5 c and in the table in figure 3.5 d, where the translation efficiencies calculated take into account the amino acid composition of these proteins with respect to frequency of serine residues found in each. Relative to the clone pIBI-MET (translation efficiency is designated as 1.0) the translation efficiencies of transcripts derived from pGEM-cysE and pGEM-CYS1 are 2.69 and 4.02 respectively. This reveals that the cysE protein is translated at an efficiency only slightly higher than that of the MET25 protein (i.e. 2.69 vs. 2.48) when the modified MET25 transcript from pSPT-METΔ-4 is translated in the RRL. In real terms, taking into account this small difference and also the sizes of the respective enzyme monomers formed, this would mean that approximately 1.76-fold more SAT enzyme monomers (273 amino acids) would be synthesized than those of OAS-OAH sulphydrylase (444 amino acids). The results of the cysE translations are particularly interesting because of the enzyme activity which has been measured in extracts from similar RRL translation experiments performed using synthetic cysE transcripts. In this case, the production of quantities of the cysE protein, serine acetyltransferase or SAT, which are similar in amount to those produced from MET25 transcripts in the RRL, enables the formation of catalytically active enzyme molecules and the detection of SAT enzyme activity. This is significant in the sense that the protein encoded by the MET25 gene, OAS-OAH sulphydrylase, is also a tetrameric enzyme. If the association constants of the monomers of SAT and OAS-OAH sulphydrylase enzymes were similar (not yet determined for monomers of the OAS-OAH sulphydrylase of S. cerevisiae ; S. Yamagata, pers. comm.) it would be

reasonable to expect that active enzyme tetramers of the latter may also be formed when this protein is translated in the RRL. (It should be noted that no enzyme activity has been observed when the SAT molecule is dissociated into monomer subunits.)

The experiments to this point have described the establishment of an inherent difference in the translation efficiency of the transcripts of the cysE, CYS1 and MET25 genes in a mammalian environment, and also a quantitative improvement in the translation efficiency of the MET25 transcripts after the addition of the Xenopus ß-globin 5'- and 3'-untranslated region and related sequences.

3.3 Detection of activity of the MET25 gene product: development of a suitable assay.

A critical facet of assessment of the suitability of the MET25 gene as a transgene involved measurement of the activity of the protein product when it was synthesized outside the yeast environment. This depended upon development of a convenient and reliable assay method; convenient in that many samples could be assayed at one time and processed quickly, and reliable in providing a sensitive, accurate and repeatable measurement of enzyme activity present in small amounts of protein in crude extracts. These features were important to the work described because the enzyme activity present in various crude in vitro protein extracts (such as those prepared following RRL translations and transient and permanent cell transfections) and in in vivo protein extracts (from the tissues of transgenic mice and sheep) was to be measured. To satisfy these requirements, the assay method devised by Yamagata (Yamagata, 1987) for measurement of the OAS sulphydrylase activity of the bifunctional OAS-OAH sulphydrylase enzyme of S. cerevisiae (i.e. of the MET25 gene product) was modified. (It should be noted that the same assay can be used to detect activity of the OAS suphydrylase enzyme of S. typhimurium.). In the first instance, the method of detection employed in the original protocol did not lend itself to multi-sample analysis of crude protein extracts of the types mentioned, especially where small amounts of protein were involved. Secondly, the reactive nature of the end product cysteine

also necessitated changes to the method, again particularly with respect to final detection.

In brief, the assay conditions set out by Yamagata included assay of the sulphydrylase enzyme at 30⁰C in the presence of 6.25 mM OAS, 25 µM pyridoxal phosphate, 1.25 mM EDTA, 1.25 mM Tris-sulphide and 125 mM potassium phosphate buffer at pH 7.8. The need to simultaneously analyse many small crude protein extracts has brought about the use of radioactive substrate molecules and adoption of the methods of thin layer chromatography (TLC) and autoradiography, each of which facilitates greater sensitivity of detection. In developing the new assay method, a crude extract from the wild-type yeast strain GP001 (an S. cerevisiae isolate; G. Pure, unpublished) was used as a source of sulphydrylase enzyme.

In addition, as the end product cysteine is rapidly oxidised and may interact with other molecules and thus escape detection, a means of rapid "cysteine capture" has been incorporated into the assay procedure. This involves the use of the alkylating agent 4-vinylpyridine (4VP), which reacts extremely rapidly with free cysteine to form a covalent linkage. The use of this substance in similar circumstances has been reported in the literature (Friedman et al., 1970; Cooper and Turnell, 1987). When included in the sulphydrylase reaction mixture at the commencement of the reaction, this compound precludes oxidation of free cysteine and produces a molecule 4-pyridyl-ethyl-cysteine, referred to here as cysteine-4VP, with properties which enable it to be detected in a number of ways. For example, following alkylation using 4VP, the usual "smear" of cysteine seen in a thin layer chromatogram (due to oxidation) is resolved into a discrete "spot" which may be visualised directly in the presence of ultraviolet (U.V.) light or indirectly by autoradiography if radioactively labelled OAS is used as a substrate. Alternatively, an HPLC procedure which also relies on the U.V.-absorbing properties of specific moieties such as pyridine may be employed to detect small quantities of cysteine-4VP formed in the sulphydrylase assay.

The assay method best suited to detection of OAS sulphydrylase enzyme activity was derived through optimisation of these reaction conditions and post-reaction processing techniques. This method can be applied to a variety of crude protein extracts and is explained in Chapter 2.

The results obtained in two experiments in which the wild-type GP001 protein extract was used as an enzyme source are shown in figure 3.6 a and b. In the first case, (a), the result of a 60 minute time course reaction is shown. In the test sample (100 µg of GP001 extract in a 200 µl reaction volume) the amount of cysteine-4VP formed is seen to increase with time. When enzyme catalysis is maximal and in the linear range (as determined by densitometric analysis of the autoradiogram) conversion of OAS to cysteine at 0.33 % / minute produces 0.823 nmoles of cysteine / minute. Where one unit of enzyme activity is defined as that able to convert 1.0 µmole of OAS to cysteine / minute, the specific activity of the GP001 extract with respect to OAS sulphydrylase activity is thus 8.23 milliunits / mg (i.e. 8.23 nmoles / minute / mg). The activity of this extract will be considered further in discussion section 3.5. As expected, no cysteine-4VP was formed in the negative control. The autoradiograph in (b) shows the sulphydrylase activity present in approximately 2.0, 4.0 and 10.0 µg of protein from the crude GP001 extract. The negative control in this case contains all substrates except for sodium sulphide, while the positive control is one in which 10.0 µg of protein from a crude cell extract made from a wild-type strain of S. typhimurium, LT2, has been used as a source of sulphydrylase enzyme. When only 2.0 µg of the GP001 extract is assayed under these conditions, some cysteine-4VP is still formed, although this appears to be near the limit of detection of the assay. These examples both serve to illustrate the capacity of the assay devised to detect authentic OAS sulphydrylase activity which is present in a crude protein extract.

3.4 Production of assayable quantities of the MET25 gene product in vitro.

With a reliable assay procedure available, measurement of sulphydrylase activity in in vitro extracts was commenced. Two systems were used in an attempt to produce assayable quantities of the MET25 gene product. Sulphydrylase activity was sought in crude protein extracts from (1) RRL translations in which capped SP6 transcripts derived from the clone pSPT-METΔ-4 (see figure 3.4 b) were used, and (2) transiently transfected cell cultures into which the expression plasmids pJL-MET and pLT-MET (see figure 3.9 a and b) were placed, or

Figure 3.6 (a) : TLC fractionation of substances in OAS sulphydrylase time course reaction samples where the GP001 yeast extract was the enzyme source.

Aliquots of 10 μl were taken from a large-scale 100 μl OAS sulphydrylase reaction (100 μg of the crude protein extract made from the wild-type yeast strain GP001 was assayed in the presence of ^{14}C-OAS) at t = 0, 1, 2, 5, 10, 20, 40 and 60 minutes, processed as detailed previously and the whole of each sample loaded onto a TLC. Fractionation of reactants and products in the samples was achieved by ascending chromatography in the solvent butan-1-ol : glacial acetic acid : water (12:3:5) and the direction of migration of the solvent is indicated on the right side of the diagram (↑). The distance of migration of the compounds L-serine (serine) O-acetyl-L-serine (OAS) N-acetyl-L-serine (NAS) and 4-pyridyl-ethyl-cysteine (cys-4VP) from the origin in the TLC is shown on the left side of the figure. An autoradiogram was produced after a 24 hour exposure of Fuji RX autoradiographic film to the chromatogram.

(b) : TLC fractionation to detect the cysteine-4VP product from OAS sulphydrylase reactions in which the enzyme was limiting.

Four OAS sulphydrylase assays of protein from the GP001 extract were carried out under identical conditions and with ^{14}C-OAS as a substrate, except that one contained 10 μg of protein but none of the vital substrate sodium sulphide (labelled -) and the other three contained different amounts of protein (2, 4 and 10 μg respectively). A positive control reaction (LT2) contained protein from a crude extract of the wild-type S. typhimurium strain, LT2. After post-reaction processing, one half of each reaction sample was fractionated in the TLC as described previously. The direction of migration of the solvent is again indicated (↑). Autoradiography of the resulting chromatogram was carried out at room temperature for 48 hours.

a.

NAS-
OAS-
serine-
cys-4VP-
origin-

0 1 2 5 10 20 40 60

Time (minutes)

b.

NAS-
OAS-
serine-
cys-4VP-
origin-

–   2   4   10   LT2

GP001
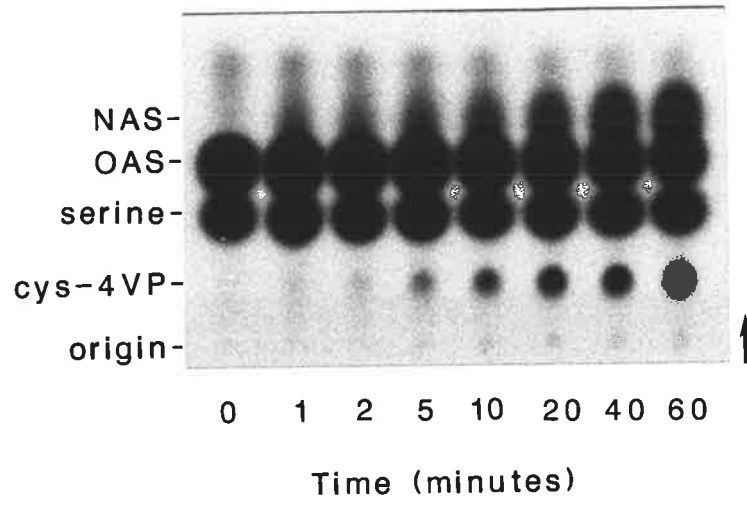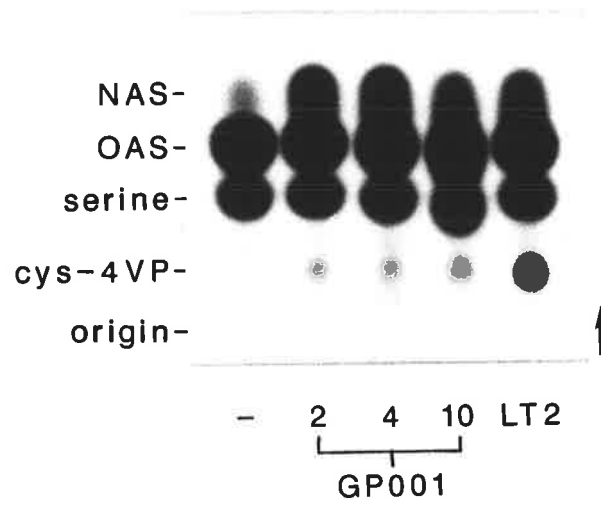
permanently transfected cell cultures where copies of the G418-selectable expression plasmid pLT-MET were integrated into the host cell genome. These in vitro systems will be explained separately and in more detail.

3.4.1 Production of assayable quantities of the MET25 gene product in the RRL.

The RRL translation system has already been considered in this chapter. Details of the optimum conditions of synthesis of the MET25 gene product have been given, and it might reasonably be expected that assayable quantities of this protein can be synthesized in the RRL, for when similar quantities of the cysE gene product of S. typhimurium are produced in the RRL, SAT enzyme activity can be detected in the translation extract. (This will be discussed in the next chapter.) With this in mind, a non-radioactive translation reaction was performed using approximately 20.0 μg/ml of pSPT-METΔ-4 transcripts and at a potassium acetate concentration of only 60.0 mM. (The use of this lower concentration, found to support near-optimal translation in the RRL (data not shown) was a modification necessary to accomodate the subsequent assay procedure.) Two control reactions done at this time included one in which transcripts from the clone pGEM-CYS1 were translated and another in which no RNA was included. Approximately 5.0 μg of protein from each of the resulting translation extracts was then assayed for sulphydrylase activity as described previously. In this experiment, and two others using lower quantities of these translation extracts, no sulphydrylase activity was detected in any of the samples. In comparison, positive control assays of GP001 (S. cerevisiae) and LT2 (S. typhimurium) extracts detected sulphydrylase activity. These results are shown in figure 3.7 a.

Regarding the lack of sulphydrylase activity, the MET25 translation extract was examined further to discover if the translation had produced assayable quantities of the MET25 gene product. This was achieved by performing a Western transfer analysis (see Methods) on an amount of the translation extracts equivalent to that used in the assays. In the same experiment, two different amounts of the positive control GP001 extract, 2.0 μg and 10.0 μg, each shown previously to contain detectable levels of sulphydrylase activity were included (figure 3.7 b). An antibody which was known to be specific for the OAS-OAH

Figure 3.7 (a) : TLC fractionation of OAS sulphydrylase assay samples in which RRL-translated protein was used as an enzyme source.

Translation of the MET25 gene product in the RRL translation system (using SP6 transcripts derived from the clone pSPT-METΔ-4, translated at the limiting potassium acetate concentration of 60 mM to avoid aberrant fractionation in the TLC) was followed by assay for OAS sulphydrylase activity (^{14}C-OAS was the substrate) and TLC fractionation as described before. One third of this translation sample was assayed alongside two negative controls (one third of translation reaction mixtures to which either no RNA or SP6 RNA derived from the clone pGEM-CYS1 had been added) and two positive controls (protein from crude extracts of the yeast strain GP001 and the bacterial strain LT2). The whole of each reaction sample was processed and fractionated via the TLC, and the resulting chromatogram autoradiographed for 72 hours at room temperature. The TLC fractionated sulphydrylase assay samples (no RNA, MET25, CYS1, GP001 and LT2) are indicated.

(b) : Western transfer analysis of crude in vitro protein extracts.

The proteins in various crude in vitro extracts were fractionated via SDS-PAGE and transferred to nitrocellulose along with molecular weight standards (BRL prestained markers) as described in the Western transfer method given in Chapter 2. The filter was then probed for the MET25 gene product (OAS-OAH sulphydrylase) using a specific antibody (Yamagata, 1984) and binding of this to target molecules visualised with the aid of a second-antibody colour detection system (Forster et al., 1985). In the photograph of this result, the respective samples are 2.0 and 10.0 μg of the yeast GP001 extract (lanes 1 and 2), 2.0 μg of HeLa cell extracts in which no DNA (lane 3) or pJL-MET DNA (lane 4) was used to transfect the cells (see section 3.4.2) and 10 μl of the RRL translation extracts described above in figure 3.7 (a) from the no RNA (lane 5), MET25 RNA (lane 6) and CYS1 RNA (lane 7) translations respectively. The position of the MET25 gene product (the 50,000 M_r monomer of the bifunctional OAS-OAH sulphydrylase of the yeast S. cerevisiae) in the Western transfer is indicated by the arrow on the right side of the figure. The sizes and migration distances of the molecular weight standard proteins (BRL prestained) is indicated on the left of the figure.

a.

No RNA | MET25 | CYS1 | GP001 | LT2

NAS—
OAS—
serine—
cys-4VP—
origin—

b.

2   10

No DNA | pJL-MET | No RNA | MET25 | CYS1

kD
97—
68—
45—

25.7—

18—

14—

◄ OAS-OAH
sulphydrylase

GP001   HeLa   RRL
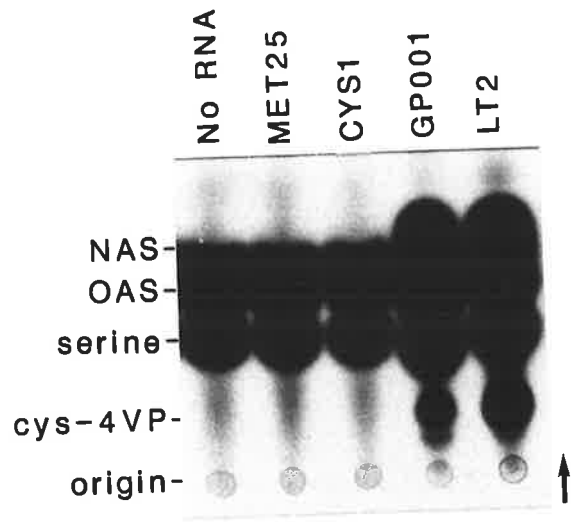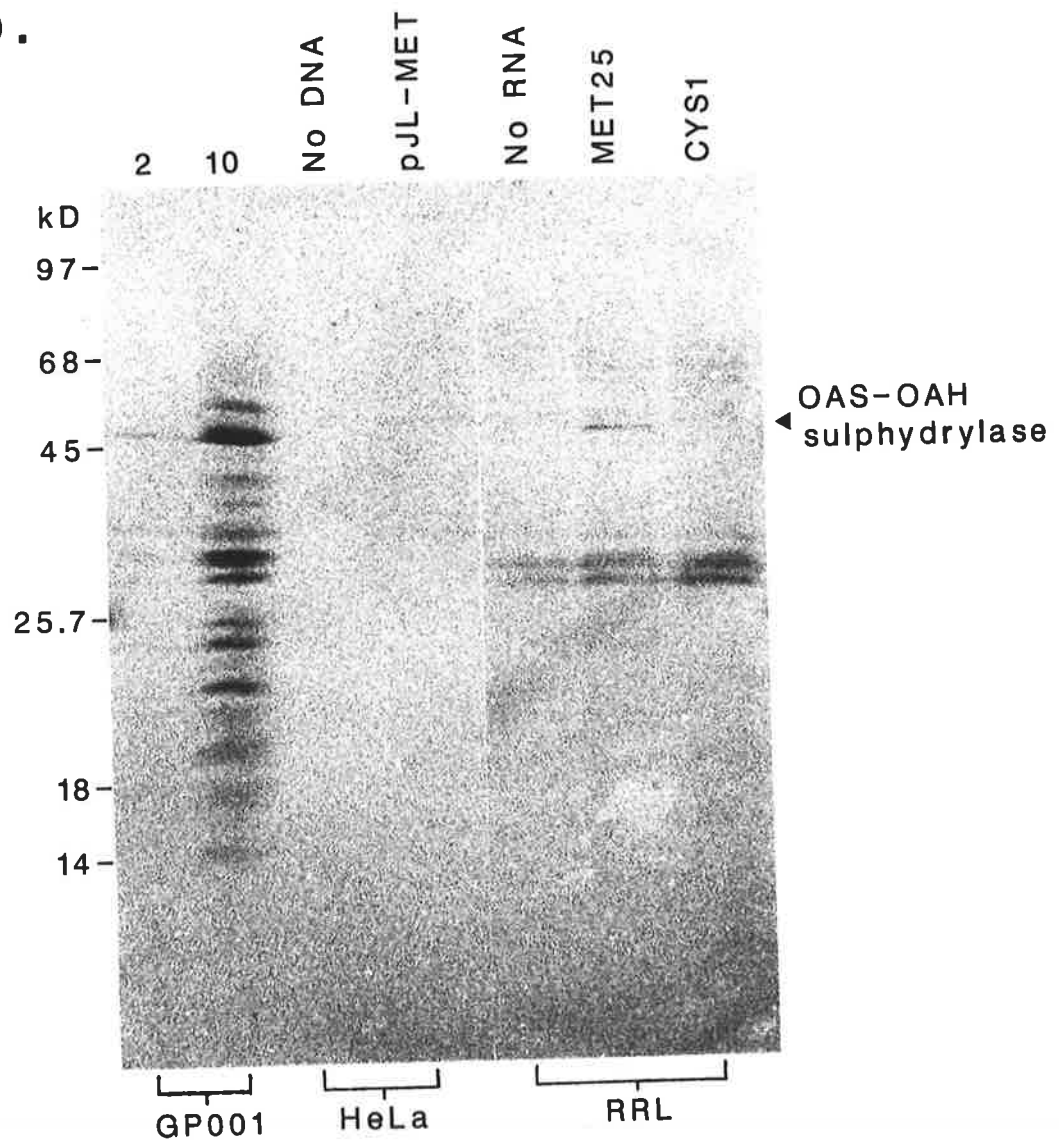
sulphydrylase enzyme of S. cerevisiae (Yamagata, 1984) was used to detect the protein. When the antiserum was diluted 5000-fold, hybridized to the Western filter and the second antibody-colour detection system applied, a single specific band of protein of the correct molecular weight for a monomer of the sulphydrylase enzyme was detected in the MET25 translation sample track (lane 6) above two non-specific bands which also appeared in the tracks of each of the negative controls (lanes 5 and 7). No specific bands were visible in either of the negative control samples. In each GP001 positive control sample, a specific band of protein also of the correct molecular weight was detected by the antibody (lanes 1 and 2). Where more GP001 protein (10.0 µg) was loaded onto the Western gel, the antibody specificity seemed lower; additional yeast proteins were detected by the antibody in this case. When comparing band intensities, it appears that an amount of the MET25 gene product slightly more than that present in 2.0 µg of the GP001 extract was present in the 5.0 µg MET25 RRL translation sample used in the unsuccessful sulphydrylase assay (figure 3.7 b).

Considering that assayable quantities of the sulphydrylase protein were detected in the RRL extract, it seemed possible that substances inhibitory to the enzyme activity could be present in the RRL. Whilst the problems caused by some ions present in the RRL have been dealt with previously, one major and essential component of this lysate which has not been considered a potential problem was haem. This is present in large quantities in the RRL in haemoglobin; a crude colourimetric determination of free and protein-bound haem in the RRL revealed a haem concentration above 2.0 mM (data not shown). This is expected considering the report of 5.75 mM haem in chicken blood (Porra and Jones, 1963). In addition to this, haem is also found in the RRL in the form of free haemin (the Fe^{2+} form of haem) which is added to an optimal working concentration of 20.0 - 40.0 µM in order to inhibit a translational repressor present in the RRL. Haem has been reported to inhibit the activity of another pyridoxal phosphate-requiring enzyme, bacterial ALA-Synthetase, apparently by way of interaction of this essential cofactor with the iron in the haem moiety found in either free haemin or in various haemoproteins such as haemoglogin. Even the ion Fe^{2+} in salts such as $FeSO_4$ was reported to cause enzyme inhibition in this case (Burnham and Lascelles, 1963).

Although the mechanism is unclear, this possibility was first tested indirectly by assaying for sulphydrylase activity in the GP001 extract in the presence of increasing amounts of added RRL solution and in the absence of added pyridoxal phosphate. When the RRL extract formed as little as 3.5% of the total volume (in a normal translation reaction extract, the RRL comprises 70% of the volume) extreme inhibition of enzyme activity was observed (figure 3.8 a, lane 1). A reaction performed under identical conditions (but with RRL added at the conclusion, just prior to TLC processing; figure 3.8 a, lane 2) demonstrated enzyme activity comparable to that seen in the positive control, GP001, which showed normal sulphydrylase activity in the absence of added RRL solution (figure 3.8 a, +). This implicated a substance(s) present in the RRL reaction mixture as being a candidate likely to be responsible for inhibition (figure 3.8 a).

A more direct determination of the involvement of haemin in inhibition of the sulphydrylase enzyme was made when pure haemin was added to the wild-type GP001 extract over a 1000-fold concentration range (2.0 µM - 2.0 mM) and the extract assayed as before, again in the absence of added pyridoxal phosphate (figure 3.8 b). At the concentration of 200 µM haemin (lane 3) marked inhibition of sulphydrylase activity was seen. At 2.0 mM haemin (lane 4) enzyme activity was not detected. In contrast, normal quantities of the cysteine-4VP product were detected in samples containing 2.0 µM and 20.0 µM added haemin (lanes 1 and 2). The possibility of interference to the TLC fractionation and subsequent detection of cysteine-4VP by increased levels of haemin was eliminated; 2.0 µM - 2.0 mM haemin added to sulphydrylase reaction mixtures at the conclusion of the reactions but prior to the TLC procedure did not appreciably affect the chromatographic separation or detection of cysteine-4VP (lanes 5 - 8).

An attempt to demonstrate restoration of sulphydrylase activity via the inclusion of increasing concentrations of pyridoxal phosphate in such reaction mixtures failed, due to the limited capacity of the TLC medium to carry this substance. Inclusion of pyridoxal phosphate at concentrations higher than 100 µM invariably caused quite aberrant results in control sulphydrylase assay samples fractionated using the TLC method (data not shown). Alternative pre-assay processing methods, involving removal of haem following translation of the protein in the RRL were not attempted.

Figure 3.8 (a) : OAS sulphydrylase assay of GP001 extract in the presence of RRL extract.

A TLC fractionation of OAS sulphydrylase reaction samples to which RRL extract was added (to 5 % (v/v) of the reaction volume) either at the commencement of the reaction (lane 1) or after the reaction had taken place (lane 2). Each reaction sample contained 10 μg of the GP001 protein extract. A control sample (labelled +) was one to which no RRL extract was added; this reaction was performed and processed for the TLC fractionation along with the other samples. One half of each sample was used in the fractionation. All assays were perfomed in the presence of ^{14}C-OAS and autoradiography was carried out at room temperature for 12 hours.

(b) : OAS sulphydrylase reactions in which GP001 extract has been assayed in the presence of increasing concentrations of haemin.

OAS sulphydrylase assays of 10 μg of the GP001 protein extract were performed in the presence of ^{14}C-OAS and increasing concentrations of haemin (2.0 μM, 20.0 μM, 200.0 μM and 2.0 mM) which was added at the commencement of the incubation period (lanes 1, 2, 3 and 4 respectively). A series of GP001 control samples (lanes 5, 6, 7 and 8 respectively) to which the same, increasing amounts of the haemin were added at the completion of the incubation period, was cofractionated in the TLC . A separate positive control GP001 assay (+) was processed and, as for the other samples, one half of this sample was fractionated in the TLC. An autoradiogram was produced after exposure of X-ray film to the chromatogram for 72 hours at room temperature.

a.

NAS-
OAS-
serine-
cys-4VP-
origin-

+ 1 2

b.

NAS-
OAS-
serine-
cys-4VP-
origin-

+ 1 2 3 4 5 6 7 8
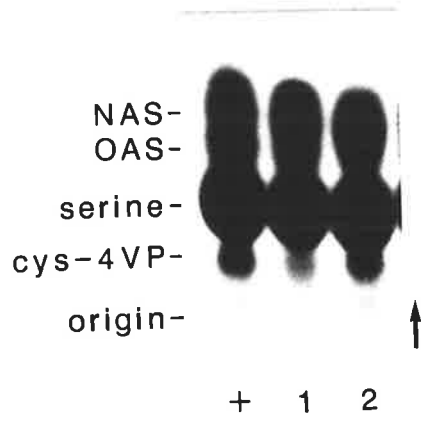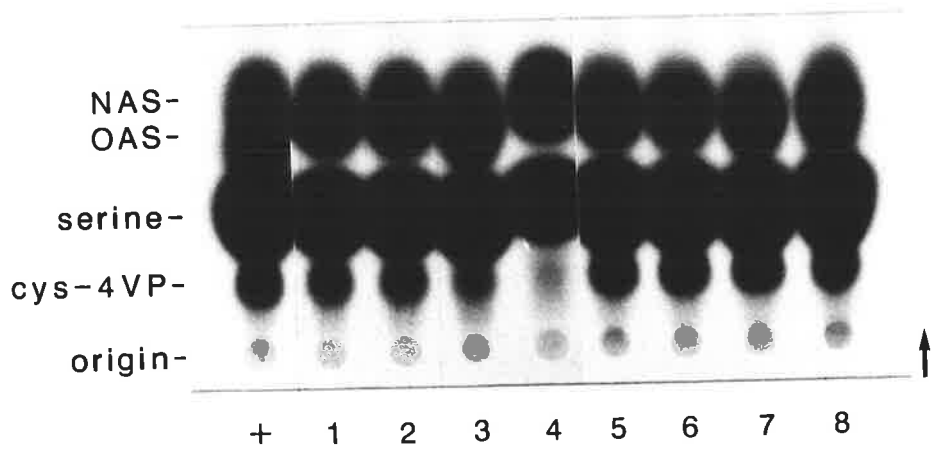
A consideration of these and other features related to translation of the MET25 gene product in the RRL (see the dicussion in section 3.5) suggested that this system was not suitable for the production of enzymically active OAS-OAH sulphydrylase. Hence, although the RRL translations had served the purpose of demonstrating improvement of translation of MET25 transcripts in a mammalian environment (and as expression of the MET25 gene was being tested by transfection of constructs into mammalian cell cultures) their use to achieve this secondary aim was discontinued. N.B. The wheat germ translation system, although eukaryotic, was not considered an appropriate alternative because it is not a mammalian system and also was assayed to have considerable OAS sulphydrylase activity of its own (data not shown).

3.4.2 Production of assayable quantities of the MET25 gene product in cell culture.

The possibility of producing active OAS sulphydrylase in transiently or permanently transfected cell culture systems was investigated in parallel with the RRL translation work. Two eukaryotic expression plasmids were constructed for use in this work. These were (a) pJL-MET and (b) pLT-MET, and are shown in figure 3.9, along with the details of their construction. The features of these constructs which are important for expression of the MET25 gene in transfected cell cultures are given here.

(a) The construct pJL-MET (figure 3.9 a) contains the entire MET25 coding region in a fragment subcloned from pSPT-METΔ-4 that includes the added Xenopus ß-globin 5' and 3'untranslated sequences shown to increase translation efficiency. This fragment is flanked at the 5' end by the SV40 late promoter and at the 3' end by the SV40 early polyadenylation and associated processing signals. These SV40 sequences have been demonstrated to facilitate expression of many different cloned genes, both from prokaryotic and eukaryotic sources, in a eukaryotic environment in vitro and in vivo (e.g. Southern and Berg, 1982; Nordstrom et al., 1985). In addition, the base vector, pJL4/A+, modified from the original vector pJL4 (Gough et al., 1985) contains the entire large T-antigen coding region and associated early promoter including the origin of replication (ori), and polyadenylation signal and processing elements, all from

Figure 3.9 (a) : Construction of eukaryotic expression vector pJL-MET.

The construction of the plasmid pJL-MET was performed in two stages. First, a 156 bp BamHI / EcoRI fragment from the vector pJL-E (see figure 4.2 a) containing the SV40 early polyadenylation signal was subcloned into the base vector pJL4 (Gough et al., 1985) which had been prepared by digestion with BamHI and EcoRI and treatment with CIP (Boehringer). This vector, named pJL4/A+ (i) was then digested with SalI, treated with KF and CIP and used as the base vector into which the MET25 gene was cloned. The MET25 gene fragment was isolated as a 1485 bp (HindIII / HinfI) blunt-ended fragment from the vector pSPT-METΔ-4 (ii) obtained by digestion with PstI, partial digestion with HinfI, partial digestion with HindIII and treatment with KF. The fragment contained the 43 bp Xenopus β-globin 5'-untranslated region bounded by 20 bp of polylinker sequences at its 5'end, the whole of the MET25 coding region including 4 bp 5' to the ATG start codon and 52 bp 3' to the TGA codon, and 36 bp of the Xenopus β-globin 3'-untranslated region. Ligation of the vector and insert produced the new plasmid construct, pJL-MET (iii). The L, ori, E and T-antigen sequences (in the SalI- AvaI/BamHI portion of this plasmid) are derived from the SV40 genome.

a.
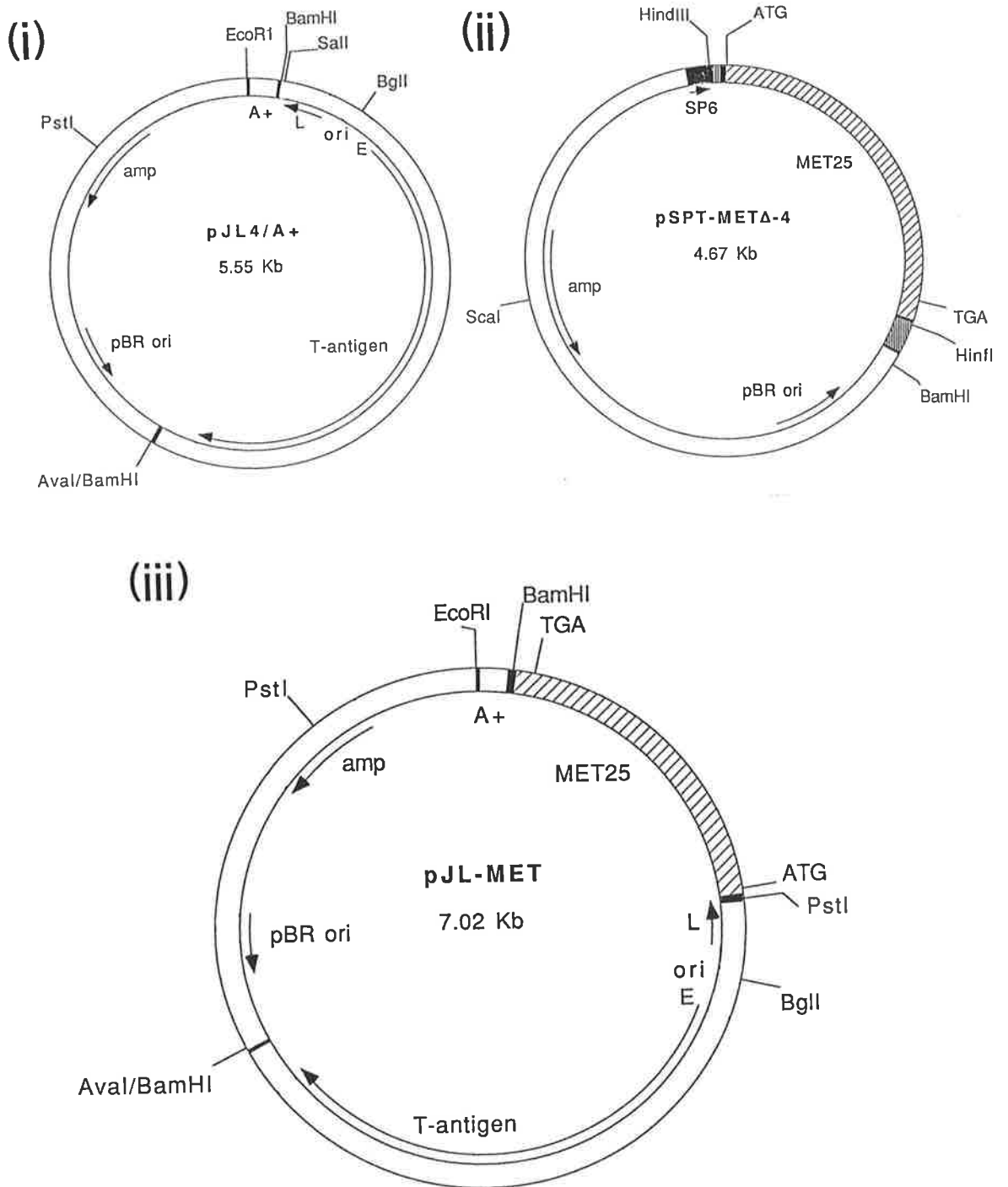
the SV40 genome. The ability of this construct to express T-antigen allows replication of the plasmid from the ori in cells which are permissive for SV40 replication (e.g. HeLa cells, African Green Monkey Kidney or CV-1 cells, or the SV40-transformed CV-1 derivative, COS-1 cells). As a multicopy plasmid, higher levels of expression of cloned genes are possible compared to similar plasmids which cannot replicate in the host cells into which they are transfected (Mellon et al., 1981; Humphries et al., 1982).

(b) The construct pLT-MET (figure 3.9 b) contains the same MET25 fragment as described above in (a), including the sequences added for improved translation, but is flanked by the RSVLTR promoter at the 5' end (Yamamoto et al., 1980) and the human Growth Hormone (hGH) polyadenylation and associated processing signals at the 3' end. These sequences are also well characterised in terms of their ability to facilitate expression of cloned genes in vitro and in vivo (Luciw et al., 1983; Overbeek et al., 1986; Pfarr et al., 1986). With a view to subsequent transgenesis experiments, the LTR promoter could serve as a useful starting point, especially in mice, where use of this promoter to direct tissue-specific gene expression has been well characterised (Overbeek et al., 1986) In addition, the gene which encodes an aminoglycoside phosphotransferase enzyme from transposon Tn5 (Rothstein et al., 1980; Auerswald et al., 1980) and confers resistance to aminoglycoside antibiotics such as neomycin or their analogues (e.g. G418) is present in pLT-MET. This gene is expressed with the aid of the SV40 early promoter and the polyadenylation and splicing signals (i.e. small t-antigen intron and splice junctions) as found in other eukaryotic expression vectors such as pSV2-CAT (Gorman et al., 1982). This feature of pLT-MET was used in creating permanently tranfected cell cultures, by selection with G418. In transient transfections, the presence of the SV40 ori region in this plasmid would enable it to replicate in cell lines (e.g. COS-1) which express T-antigen constitutively.

Transient transfection of HeLa, Chinese Ham ster Ovary (CHO) and COS-1 cell lines was carried out using each of these constructs according to the methods described in Chapter 2. A negative control in which no DNA was transfected into the cells, was included in the case of each different cell line, as were positive controls in which DNA from either of the similarly constructed

Figure 3.9 (b) : Construction of eukaryotic expression vector pLT-MET.

To construct pLT-MET, a 1505 b.p. PstI / BamHI digested and KF treated fragment was isolated from the plasmid pJL-MET (i) and subcloned into the base vector pRSVΔH (ii) which was prepared by BamHI digestion and treatment with KF and CIP. The insert fragment was identical to that in pJL-MET and contained the MET25 gene flanked by the sequence elements described in the legend to figure 3.9 (a). In the new plasmid pLT-MET (iii) the insert itself is flanked by the promoter contained in the Rous Sarcoma Virus Long Terminal Repeat region (RSVLTR; 535 bp) at its 5' end and the human Growth Hormone polyadenylation 3' processing sequences (622 bp) at its 3' end. This plasmid also contains the aminoglycoside phosphotransferase (neo^r) gene, which when expressed, confers resistance on the host cell, to the aminoglycoside antibiotics such as neomycin, or its analogue G418. This gene is expressed in eukaryotic cells via operation of the SV40 early promoter. The unique SalI restriction site in this plasmid does not impinge on any functional elements of either the RSVLTR or SV40 early promoter. This is a site which was convenient for use in linearisation of the plasmid prior to the permanent CHO cell transfection experiment. In the linear map in (iv) the position of a few key restriction enzyme sites (relevant to probe fragment used, and to sizes of restriction fragments expected) is given.

b.

(i) pJL-MET 7.02 Kb — EcoRI, BamHI, TGA, PstI, amp, A+, MET25, pBR ori, ATG, PstI, L, ori, E, BgII, Aval/BamHI, T-antigen

(ii) pRSVΔH 6.51 Kb — SalI, EcoRI, BamHI, RSVLTR, hGH, SV40, amp, ScaI, neo, pBR ori

(iii) pLT-MET 8.00 Kb — HindIII, ATG, EcoRI, EcoRI, SalI, RSVLTR, MET25, HindIII, SV40, TGA, neo, hGH, pBR ori, amp, ScaI

(iv) SalI, EcoRI, HindIII, EcoRI, HindIII, ScaI, 444 b, 91 b, 460 b, 705 b, 1827 b, RSVLTR, MET25, hGH
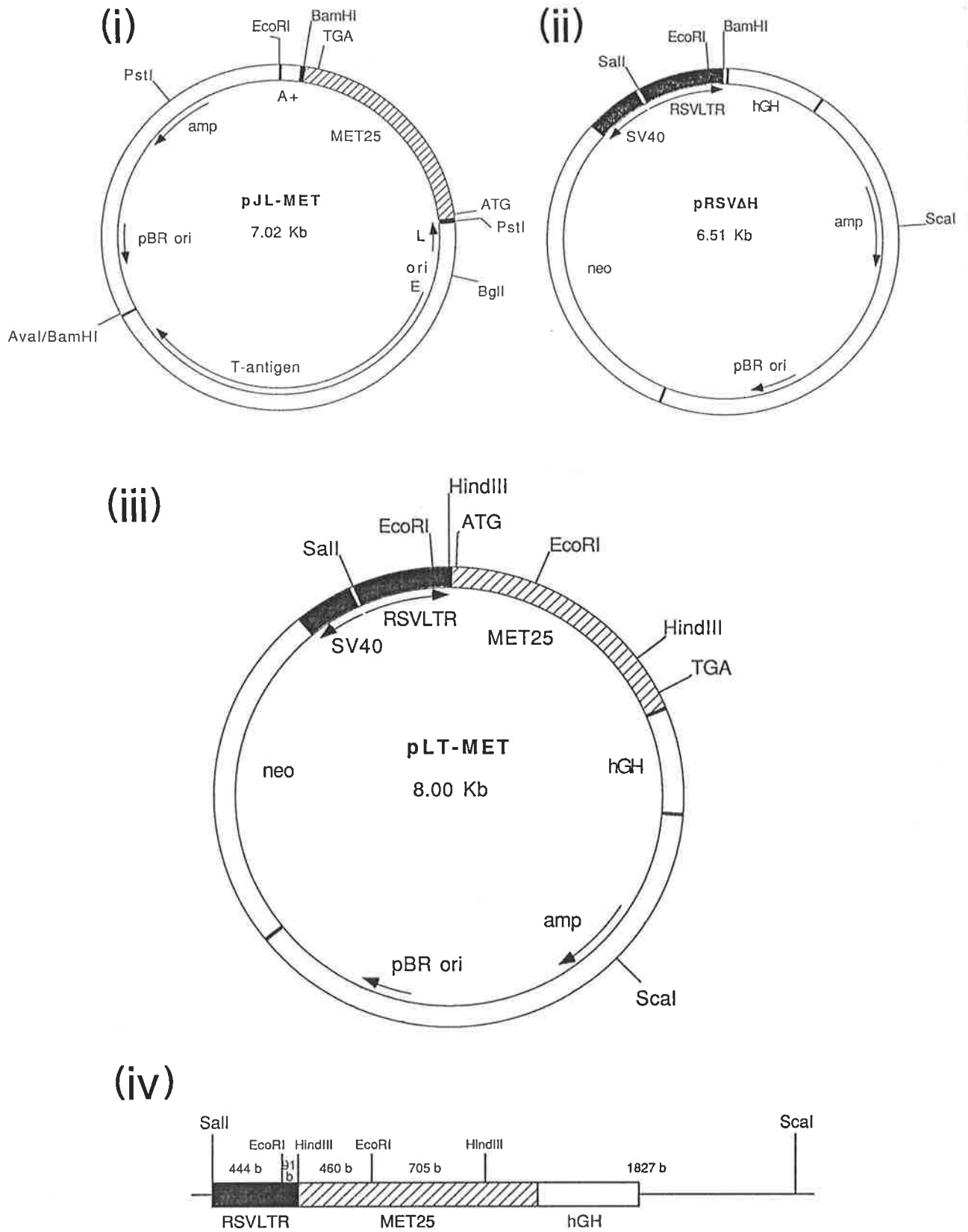
cysE gene constructs pJL-E and pLT-E (see figure 4.2) were tranfected. In repeated assays using portions of the crude protein extracts made from cells harvested at the conclusion of each of these transfection experiments, no sulphydrylase activity was detected. In addition, a primer extension analysis of total RNA prepared from HeLa cells transfected with pJL-MET failed to reveal detectable levels of MET25 transcripts (data not shown). In the case of the cysE gene positive controls, protein extracts made from the transfected cells invariably exhibited the expected SAT activity when assayed. The results of the CHO and COS-1 transfections using the positive control cysE constructs are shown in figure 4.3. The presence of cysE transcripts in total RNA prepared from the transfected HeLa cells and analysed by the method of primer extension supported the SAT assay result (data not shown).Details of the cysE gene expression work will be given in the next chapter).

A Western transfer experiment was also performed to ascertain production of MET25 protein (OAS-OAH sulphydrylase) in the transfected cells. The result of this experiment is shown in section 3.4.1 (see figure 3.7 b) where approximately 2.0 μg of protein from the HeLa cell extracts of the transfections containing no DNA and pJL-MET DNA (tracks 4 and 5) have been probed with the same OAS-OAH sulphydrylase antibody used to detect MET25 protein in the GP001 and RRL translation extracts (discussed previously in 3.4.1 and seen in lanes 1 and 2, and 5, 6 and 7 of figure 3.7 b. Clearly, no OAS-OAH sulphydrylase is detected by the antibody in the HeLa cell extract from either the pJL-MET or negative control (no DNA) transfection.

In a final attempt to produce assayable quantities of the MET25 gene product in vitro, the expression plasmid pLT-MET was used to transfect CHO cells in order to create permanent cell lines which contain intact and functional copies of the pLT-MET construct, i.e. of both the neo[r] and MET25 genes. To this end, pLT-MET DNA was linearised with the restriction endonuclease SalI, which cleaves the DNA at the junction point between the divergent RSV-LTR and SV40 promoters of the respective MET25 and neo[r] genes in this construct (see figure 3.9 b). Following this, the DNA was purified by spermine precipitation and transfected into CHO cells by the electroporation method (see Chapter 2). After two weeks of selection of transfectants in medium containing 350 μg/ml G418,

ten G418-resistant clones were isolated (pLT-MET.1 - pLT-MET.10). The negative control for this experiment was the base vector pRSVΔH (see figure 3.9 b showing the construction of pLT-MET), which itself was linearised with SalI restriction endonuclease and transfected into CHO cells in the same manner; two pRSVΔH cell clones were isolated following similar G418 selection. These clones, pRSV.1 and pRSV.2, are G418-resistant but cannot synthesize the sulphydrylase enzyme.

The cloned cell lines isolated were kept in culture for 3 months and each retained its G418-resistant character, indicating the permanence of the integrated plasmid DNA in the CHO cell genome. The presence of integrated plasmid sequences in the genomes of the ten pLT-MET cell lines and the two pRSVΔH cell lines was established by dot blot hybridisation analysis of genomic DNA prepared from each cell line (figure 3.10 a). The general probe used in this analysis was the 3.6 kb SalI / ScaI fragment from pLT-MET, which contained the entire MET25 coding region and the 5' RSVLTR and 3' hGH processing sequences. This probe could also detect sequences in the base vector pRSVΔH. When radiolabelled by Nick-translation using ^{32}P-dNTP's and hybridised to these DNA samples, this fragment revealed low numbers of integrated plasmids (1 to 5 copies) in the genomes of the pLT-MET and pRSVΔH cell lines, with the largest number (5 copies) present in the genomes of lines pLT-MET.2, pLT-MET.6 and pLT-MET.9. When this dot blot was stripped and reprobed with an 1165 bp HindIII / HindIII MET25-specific coding region probe, a similar hybridisation pattern was observed for the ten pLT-MET cell line DNA samples, but the probe did not hybridise to the pRSV.1 and pRSV.2 DNA samples (data not shown).

A Western transfer analysis of 5.0 μg of the protein from a crude cell extract made from each cell line was carried out using the method and antibody previously described. The positive control was 5.0 μg of protein from the GP001 extract and the negative control 5.0 μg of protein from the cell line pRSV.1. The result showed that the OAS-OAH sulphydrylase-specific antibody detected a protein species of the correct molecular weight in the extracts of only two of the cell lines, pLT-MET.2 and pLT-MET.9, but not in the negative control extract from pRSV.1 nor in the extracts made from any of the other pLT-MET cell lines

(figure 3.10 b). The OAS-OAH sulphydrylase band was also detected in the positive control sample GP001. One other non-specific protein band was visible in all the cell line extract samples including the negative control. Comparing the intensities of the specific bands detected in the extracts made from cell lines pLT-MET.2 and pLT-MET.9 with that of the GP001 sample in this Western and that shown in figure 3.7 b, it could be estimated that these cell lines produce assayable quantities of the sulphydrylase enzyme. The Western transfer result for GP001, pRSV.1, pLT-MET.2 and pLT-MET.9 is shown in figure 3.10 b.

The immunological detection of sulphydrylase enzyme in the crude protein extracts of cell lines pLT-MET.2 and pLT-MET.9 led to immediate testing of these extracts for sulphydrylase activity (figure 3.10 c). In repeated assays, no cysteine-4VP was detected in either of the test samples; as in the negative control, only the unreacted substrate ^{14}C-OAS was detected in reaction mixtures at the conclusion of the incubation period. In contrast, as evidenced by the cysteine-4VP formed, the positive control extract GP001 still exhibited sulphydrylase activity. Evidently, even when assayable quantities of the MET25 gene product, OAS-OAH sulphydrylase, are synthesized in a mammalian cell culture system and assayed under the same conditions established to detect sulphydrylase activity in the GP001 extract, no enzyme activity is detected.

Further characterisation of cell lines pLT-MET.2 and pLT-MET.9 included a Southern transfer analysis, performed using genomic DNA prepared from these cells and from the negative control cell line pRSV.1. Detection of appropriately sized fragments in the EcoRI and HindIII digested genomic DNA (7.45 kb and 0.55 kb for EcoRI, and 6.84 kb and 1.17 kb for HindIII) indicated multiple intact copies of the construct inserted in the genomes of the pLT-MET cell lines in a head to tail array (figure 3.10 d). The sizes of all main bands visible in this autoradiogram (4 main bands seen in each EcoRI track and 5 main bands in each HindIII track in this, and in repeated Southern analyses) are tabulated in figure 3.10 e. On the basis of the different sizes of the 5' and 3' end-fragments associated with genomic DNA flanking the integration locus, the cell lines were judged to be different cell clones.

Figure 3.10 (a) : Dot Blot gene-copy analysis of genomic DNA from G418-resistant pLT-MET CHO cell lines.

Genomic DNA was prepared from cells of the G418-resistant pLT-MET cell lines and from lines pRSV.1 and pRSV.2 and 5 μg transferred to Genescreen membrane (Du Pont) using a dot blot apparatus (Schleicher and Schuell) and according to the manufacturers' specifications. Appropriate amounts of pLT-MET plasmid DNA were combined with 5 μg of negative control mouse genomic DNA in order to make a copy control series, and these also transferred to the Genescreen membrane. No DNA was loaded onto the membrane in areas marked (-). A MET25 gene probe (the entire 3.6 kb SalI / ScaI fragment from expression clone pLT-MET) was prepared and radioactively labelled by Nick-translation in the presence of ^{32}P-dATP/dCTP (Bresatec) then used to probe the membrane for MET25 sequences. Following filter washing (for 2 hours at 65 ^{0}C in 0.2 x SSC) autoradiography was carried out at -80 ^{0}C for 20 hours.

(b) : Western transfer analysis of protein extracts made from G418-resistant pLT-MET CHO cell lines.

Protein extracts were made from subconfluent monolayer cultures of the G418-resistant cell lines pRSV.1, pLT-MET.2 and pLT-MET.9. Approximately 5 μg of protein from each was fractionated via SDS-PAGE (12.5 % gel) transferred to nitrocellulose and probed with the antibody specific for the MET25 gene product OAS-OAH sulphydrylase. A positive control (5 μg of protein from the GP001 extract) was included in this analysis. Molecular weight standards cofractionated with the experimental samples were used to confirm the size of the species detected by the antibody; the size (kD) and migration distance is shown on the left of the figure. The position of the MET25 gene product detected is indicated on the right of the figure (◄).

(c) : OAS sulphydrylase assays of protein from the G418-resistant cell lines pRSV.1, pLT-MET.2 and pLT-MET.9.

The protein extracts used in the Western analysis described in figure 3.10 (b) were assayed for OAS sulphydrylase activity in the presence of the substrate ^{14}C-OAS; 5 μg of protein from CHO cell lines pRSV.1, pLT-MET.2 and pLT-MET.9 were assayed alongside 5 μg of the GP001 extract. After processing, one half of each reaction sample was fractionated in the TLC, along with the ^{14}C-labelled standards, serine, OAS and NAS. Autoradiography was carried out at room temperature for 48 hours.

a.

pLT–MET clones

pRSV clones

.1 .2 .3 .4 .5 .6 .7 .8 .9 .10 .1 .2

0 0.5 1 2 5 10 25 50 100

Copy no. controls

b.

GP001   pRSV.1   pLT–MET.2   pLT–MET.9

kD

45–   ◄ OAS–OAH sulphydrylase

25.7–

c.

NAS–
OAS–
serine–

cys–4VP–

origin–

^{14}C–SER ^{14}C–OAS ^{14}C–NAS   GP001   pRSV.1   pLT–MET.2   pLT–MET.9
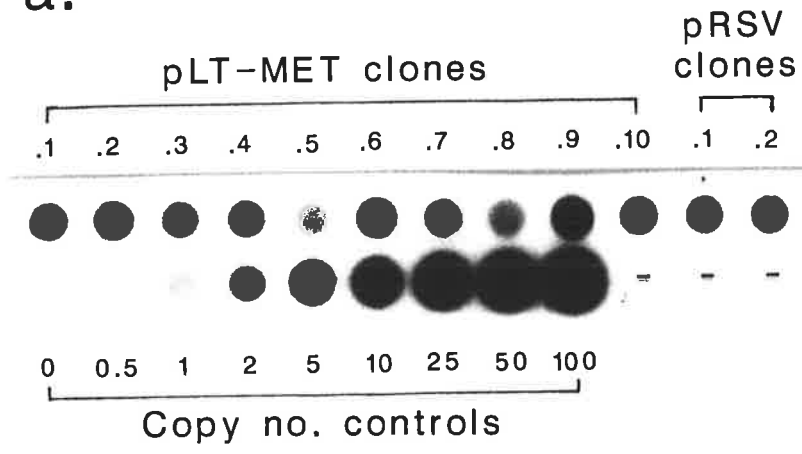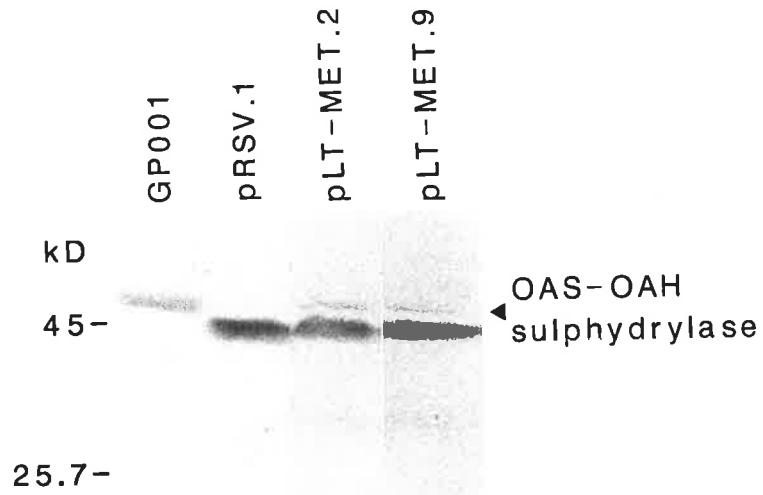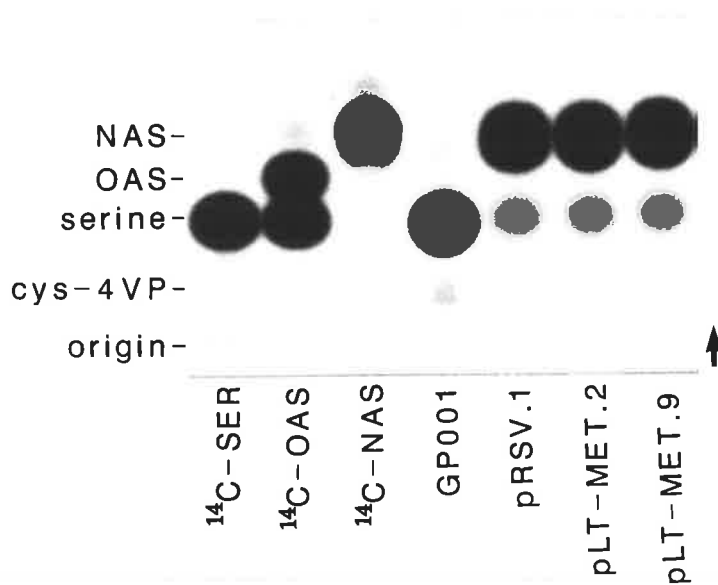
Figure 3.10 (d) : Southern transfer analysis of EcoRI and HindIII digested genomic DNA from cell lines pLT-MET.2 and pLT-MET.9.

Genomic DNA from the cell lines pLT-MET.2 and pLT-MET.9 was digested with restriction endonucleases EcoRI and HindIII and 5 µg of DNA from the resulting digests fractionated by electrophoresis in 1 % (TAE) agarose, transferred to Zetaprobe membrane and the transferred DNA probed with the 3.6 kb SalI / ScaI Nick-translated ^{32}P-labelled probe, as used in the dot blot. DNA from the cell line pRSV.1 was digested with EcoRI and analysed in the same manner, as a negative control. The filter was washed at 65 ^{0}C and 0.2 x SSC for 2 hours prior to autoradiography (16 hours, -80 ^{0}C).

(e) : Sizes of DNA fragments detected by the SalI / ScaI probe used in the Southern analysis of pLT-MET cell lines.

In the table is a list of the sizes of the bands visible following the autoradiography described in (d) above. The presence of common bands and bands unique to the separate cloned lines pLT-MET.2 and pLT-MET.9 is indicated (◄).

d.
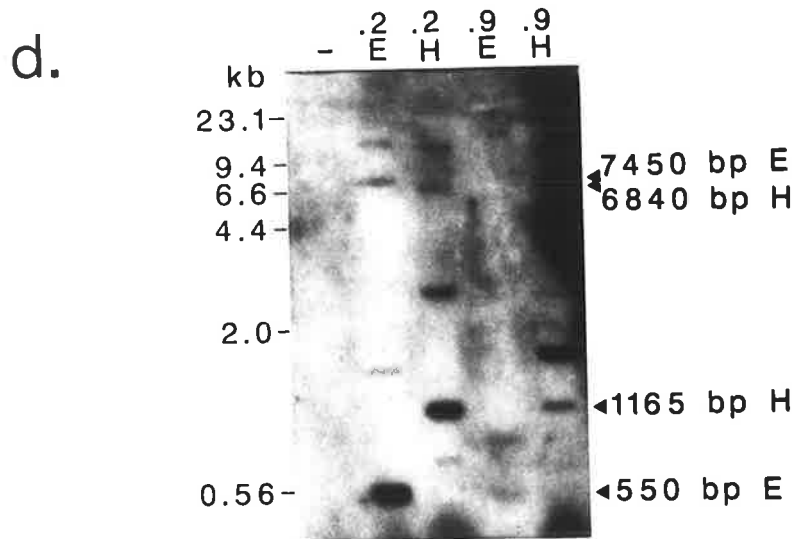
e.

Approx. sizes of hybridisation bands (in kb).	pLT-MET.2 EcoRI	pLT-MET.2 HindIII	pLT-MET.9 EcoRI	pLT-MET.9 HindIII
Common bands	7.45	6.84	7.45	6.84
	0.55	1.17	0.55	1.17
Other bands	16.2	13.8	19.6	11.2
	1.5	2.7	0.9	1.6
	-	0.8	-	0.7

Discussion.

3.5 The suitability of MET25 as a transgene.

In order to assess the suitability of the MET25 gene of S. cerevisiae for use in the current sheep transgenesis programme, a series of experiments has been completed which relate to the expression of this gene in a mammalian environment. These included analysis and improvement of the translation efficiency of MET25 transcripts in a mammalian in vitro translation system, development of an assay procedure to detect OAS sulphydrylase enzyme activity in small samples of protein in crude extracts, and finally production of assayable quantities of the sulphydrylase enzyme by expression of the MET25 gene in cultured cells.

3.5.1 Determination of translation efficiency.

In the first of these experiments, it was demonstrated that the translation efficiency of MET25 gene transcripts in a mammalian environment, namely the RRL, was quite poor when compared to the transcripts of the cysE and CYS1 genes. A translation efficiency ratio of 1.0 : 2.7 : 4.0 was measured for the translation of equal concentrations of capped SP6 transcripts derived from the constructs pIBI-MET, pGEM-cysE and pGEM-CYS1 respectively (see the table in figure 3.5 d) which took into account amino acid composition with respect to serine content (^{3}H-serine was used in the translations), but was calculated irrespective of the different sizes of the protein products (i.e. considering the rates of initiation and elongation of translation to be the same in the case of each protein). Three factors could interfere with the parameters of translation just mentioned, namely secondary and tertiary structure in the mRNA (particularly at the 5' end, in the untranslated region), poor codon usage (with respect to codons most frequently used in mammalian translation) and a poor initiation context (with respect to the sequence immediately surrounding the initiation codon).

An analysis of the MET25 SP6 transcript (derived from pIBI-MET for use in the RRL translations; previously described in section 3.2.1) and also the cysE and CYS1 SP6 transcripts with respect to secondary structure formation was performed utilising the "Fold" programme devised by Zuker and Stiegler (found

in the University of Wisconsin Genetics Computing Group (GCG) Sequence Analysis software package and described by Devereaux et al., 1984). The portions of sequence analysed in this fashion included the entire 5'-untranslated region and first 6 codons of each of the transcripts mentioned above. These are shown in figure 3.11 and the position of the ATG codon in each folded transcript is indicated. The programme, which finds an optimal secondary structure of minimum free energy in an RNA molecule, predicted the possibility of stem-loop structures (which may hinder processing by the translational machinery) in the 5'-untranslated region of the MET25 gene transcript and in those of cysE and CYS1, including the sequences contributed to the 5' terminus of each transcript by the SP6 promoter sequences (figure 3.11). The secondary structures predicted vary in complexity, size and in the minimum free energy associated with their formation.

When compared with the result of a Fold analysis of the well documented 5'-untranslated region of ferritin mRNA (ΔG^0 = -15.1 kcal/mole; Wang et al., 1990) the value of the minimum free energy of formation of the structures analysed here (ΔG^0 = -9.2, -34.0 and -10.4 kcal/mole for MET25, cysE and CYS1 transcripts respectively) suggests that each of the three transcripts may form such a secondary structure in vitro. According to the literature, one would expect more efficient translation of a transcript with minimal secondary structure in the 5'-untranslated region. As the initiation codons in the CYS1 and MET25 transcripts are found outside the regions of secondary structure predicted for these RNA molecules (the CYS1 transcript features a large unstructured region preceding the AUG codon) these transcripts may be expected to be translated more efficiently than that of cysE (according to the secondary structure prediction for this RNA, the AUG triplet is contained in a long stem). This provides some consistency considering the relatively high translation efficiency of the CYS1 transcript in the RRL (4.0), but not if considering the lower efficiency of translation of pIBI-MET transcripts (1.0) or the relatively efficient translation of cysE transcripts (2.7). This is especially disparate considering that cysE is the most likely of these three transcripts to form a secondary structure.

Without further consideration of the result of this Fold analysis, and in the absence of information regarding (a) the contribution to translation efficiency

Figure 3.11 : Comparison of secondary structures predicted in the 5'-untranslated regions of the SP6 transcripts produced from the transcription clones pIBI-MET, pGEM-cysE and pGEM-CYS1.

The figure shows the result of a "Fold" analysis performed using the programme of Zuker and Stiegler (1981) which predicts the structures (with minimum free energies of formation) that are likely to form in RNA molecules in solution. In each case, the analysis has been performed on the sequence contained in the 5'-untranslated region and including bases up to those in codon 6 of the coding region of the synthetic RNA transcripts derived from these clones. In the figure, the 5' and 3' termini of these segments are indicated, as is the position of the A residue of the ATG codon in each transcript. The Gibb's free energy of formation (-kcal/mol) and the length (nucleotides) is also given for each transcript.

a.   pIBI-MET   Length: 136 Energy: -9.2
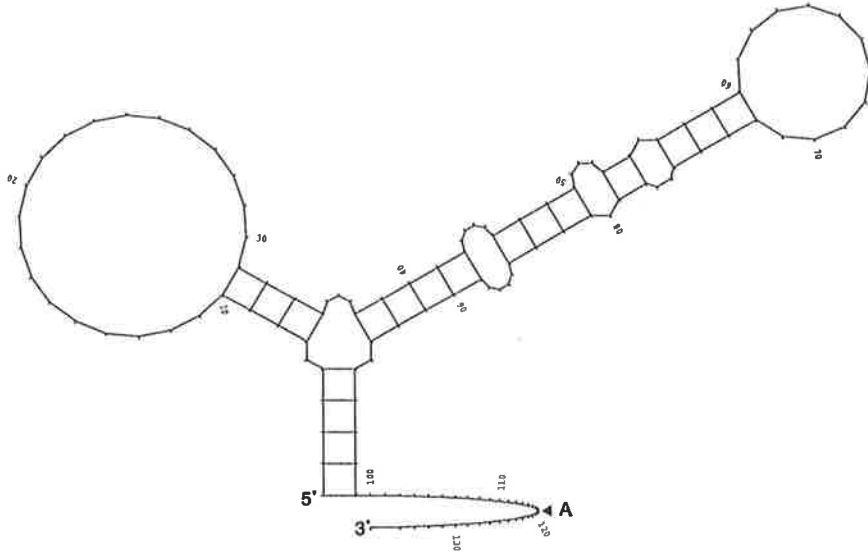
b.   pGEM-cysE   Length: 132 Energy: -34.0
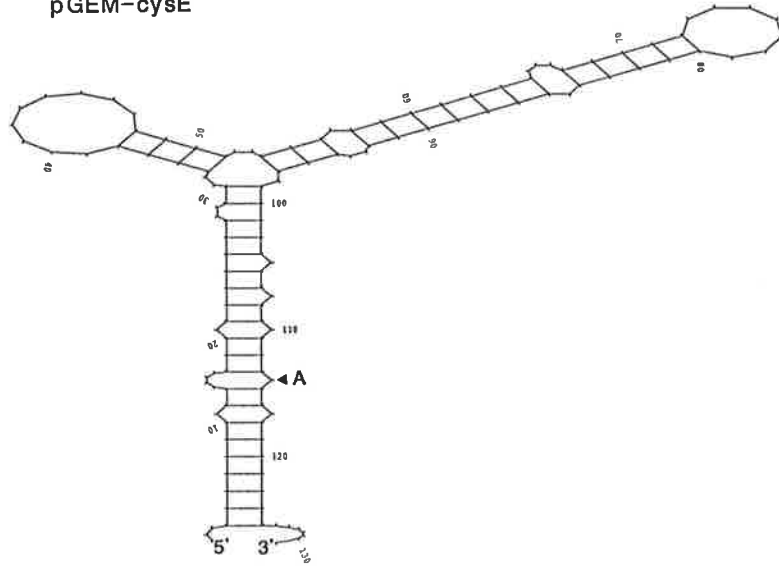
c.   pGEM-CYS1   Length: 124 Energy: -10.4
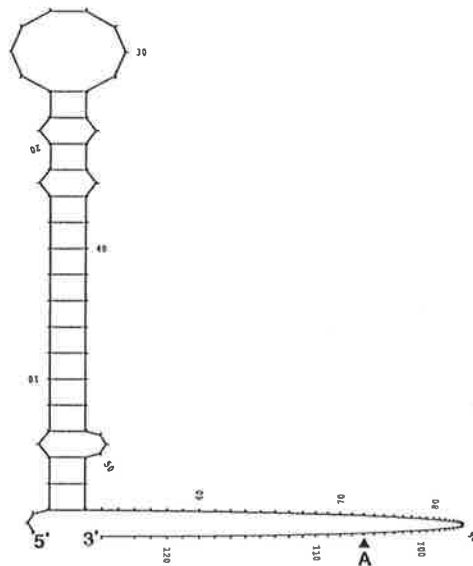
afforded by the 3-untranslated regions of these particular transcripts or (b) specific interactions which may occur between these RNA molecules and ribosomes and associated translation factors in the RRL, it is accepted that an extremely complex analysis would be required to determine the way in which differing degrees of secondary and tertiary structure might contribute to differences in the efficiency of translation of these three RNA species. Even then, such analytical predictions may not give an accurate measure of the behaviour of an mRNA molecule in a translation reaction.

Codon usage in the MET25, cysE and CYS1 transcripts was analysed with respect to the more frequently used nucleotide combinations in the codons of mammalian genes (and hence availability of certain tRNA species). The "ANALYSEQ" programme devised by Staden (Staden, 1984) was used in this analysis. This revealed no obvious disadvantage in the case of any particular gene. Codons quite infrequently used in mammalian genes are found very early in the 5' end of trancripts of CYS1 and MET25 alike, and also at other points in the sequence. For example, three leucine codons in MET25 (codons 10, 49 and 74) are the rarely used forms CTA and TTA (combined frequency of use of these leucine codons in rabbit is only 2.5 % of all leucine codons, and in eukaryotes in general is only 10.2 %. In yeast, the combined frequency of use of these two leucine codons is much higher, at 35.3 %; see the codon usage comparisons made by Aota et al., 1988). In CYS1, codons 3, 68 and 132 are also these rarely used leucine codons. While the availability of certain tRNA species in a mammalian cell may place an upper limit on the level of translation of such "yeast-specific" codons, CYS1, unlike MET25, is translated well.

An examination of the sequences which create the environment for the initiation of translation of these three gene transcripts (i.e. those in and around the initiation codon) shows that the cysE transcript is at the greatest disadvantage, if adhering to the rules of Kozak's original consensus AXX<u>AUG</u>G (where X may be any of the four nucleotides; Kozak,1984). In this respect however, the MET25 gene transcript provides a near ideal initiation sequence configuration, one which is shared by a number of well expressed genes from many different species (Kozak,1981) including such sequences as goat ß-globin, TMV coat-protein and the <u>Drosophila</u> actin and 70K heat shock proteins. More recently, extended

consensus nucleotide configurations for the translation initiation environment have been proposed (Cavener, 1987; Kozak, 1987). Analysis of the 5'-untranslated regions of the cysE, CYS1 and MET25 transcripts in relation to the more recent consensus proposed by Kozak (GCCGCCA/GCCAUGG) reveals that none of these RNA species share significant homology with the sequence added to extend the 5' end of this motif.

In addition to the above, differential stability of the transcripts in solution, due to special features of the 5'- or 3'-untranslated regions, or the coding regions themselves, may cause measured differences in translation efficiency.

3.5.2 Improvement of translation efficiency.

To circumvent lengthy analysis of the aforementioned possibilities in the case of the MET25 transcript, it was decided to remove the existing 5'-untranslated region (much of the 3'-untranslated region of the MET25 gene had been removed in the construction of the initial transcription vector pIBI-MET) and replace these regions with heterologous sequences known to confer greater stability and translation efficiency on trancripts in eukaryotic environments. From amongst the few reported in the literature, and available in the laboratory at the time, the Xenopus ß-globin sequences in the base vector pSP64-T (Krieg and Melton, 1984b) were chosen to improve the translation efficiency of MET25 transcripts. These sequences have been shown to improve translation efficiency in the Xenopus oocyte and wheat germ lysate translation systems (i.e. they are not specific to cell type or species).

A comparison of the translation efficiency of SP6 transcripts from pIBI-MET and those of three new MET25 constructs (pIBI-METΔ-4, pSPT-MET and pSPT-METΔ-4) in the RRL translation system was made (see table 3.1). The pSPT clones are the same as the corresponding pIBI clones, except for the addition of the heterologous sequences mentioned (i.e. the 43 base Xenopus ß-globin 5'-untranslated region placed 5' to the MET25 gene, 145 bases from the 3'-untranslated region of the same globin gene placed 3' to the MET25 gene, and a 23 base poly-dA / 30 base poly-dC nucleotide stretch placed further 3' to the MET25 gene. The arrangement of these sequences with respect to the MET25 gene is shown in figure 3.4 b). Differences related to the efficiency of translation

were observed when comparing the pair of clones pSPT-MET and pSPT-METΔ-4 to each other, and when comparing these to their corresponding pIBI-based counterparts, pIBI-MET and pIBI-METΔ-4. First, the transcripts from pSPT-METΔ-4 were translated more efficiently than those from pSPT-MET. Secondly, the transcripts from either of the pSPT clones, pSPT-MET and pSPT-METΔ-4, were translated with a much higher efficiency than those from the corresponding pIBI clones, pIBI-MET and pIBI-METΔ-4. Clearly, the higher translation efficiency observed for transcripts containing the ß-globin sequences was optimal in the absence of the upstream yeast sequences.

In conflict with this is the lack of improvement in translation when these sequences were removed from pIBI-MET to construct pIBI-METΔ-4. However, although the extra 100 bp of yeast sequence 5' to the MET25 gene was not present in pIBI-METΔ-4 (c.f. pIBI-MET) the reduced size of the 5'-untranslated region (only 15 bases were present 5' to the AUG codon in these transcripts) may in turn have adversely affected a possible improvement in translation efficiency in this case.

Overall, this analysis had identified a significant improvement in the translation efficiency of the MET25 gene transcripts and provided a MET25 gene "cassette" which could be used in further expression studies. Consequently, no further plasmid constructs were designed to distinguish the relative contributions of the 5' and 3' ß-globin sequences to the observed enhancement.

3.5.3 Determination of the catalytic capacity of MET25 protein translated in the RRL.

When compared to the translation efficiency of cysE transcripts, the enhanced level of translation of MET25 transcripts in the RRL was considered sufficient to allow detection of sulphydrylase enzyme activity. Indeed, assayable quantities of the MET25 gene product were shown to be produced in the RRL by Western analysis. Despite this, sulphydrylase activity was not detected in these RRL extracts. One factor likely to be a cause of this is indirect inhibition of the enzyme by the interaction of the haem in the RRL with pyridoxal phosphate, a cofactor vital for enzyme activity. Although the mechanism of interaction is unclear, it was demonstrated that the level of haem in the RRL, found free (as

added haemin) and bound in protein, is capable of abolition of the activity of the OAS sulphydrylase activity in a yeast protein extract made from the wild-type strain GP001. Another possibility was the inhibition of enzyme activity by methionine (at 1.0 mM) in the RRL translation. Though unlikely (considering that the K_i of L-methionine for yeast OAS sulphydrylase is 3.6 mM, and that even at 30 mM L-methionine, only 73 % inhibition of activity occurs) this was tested directly using the GP001 extract. When no inhibition of sulphydrylase activity in the GP001 extract was detected using this assay method (even at 25 mM methionine; data not shown) this was ruled out as being a factor contributing to inactivity of the MET25 gene product synthesized in the RRL. Likewise, reaction of synthesized cysteine with pyridoxal 5'-phosphate itself (see section 6.2) in preference to 4-vinylpyridine was discounted as a reason for lack of detection of cysteine in the GP001 assays. Other factors possibly contributing to the lack of activity are discussed later.

3.5.4   Production of the MET25 gene product in cell culture.

As an alternative to the RRL translation system, which had been useful in analysis of translation efficiency of the MET25 gene transcript, expression of the modified MET25 gene in cell culture was attempted as it more closely resembled the mammalian in vivo situation, especially with respect to post-translational modification of the gene product. In addition, it was expected to more readily allow production of assayable quantities of the MET25 gene product. So, to provide a final assessment of the suitability of the MET25 gene for use in transgenesis, expression of the modified MET25 gene was attempted by transient transfection of the plasmid constructs pJL-MET and pLT-MET (see figure 3.9) into a number of cell lines. The SV40 and RSVLTR promoters in these constructs were chosen as being among the most active promoters of eukaryotic transcription in such mammalian cell lines. In addition, each of the 3'end polyadenylation and processing signal sequences included in the constructs have been shown to enable accurate processing of transcripts. Combined with the improved translation efficiency of MET25 transcripts, it was assumed that high levels of transcription would facilitate maximal expression of the cloned MET25 gene. However, alongside successful positive control transfections using similarly constructed

cysE gene plasmids, no sulphydrylase activity, nor assayable quantities of the MET25 gene product (as judged by Western analysis) were detected. Lack of detectable levels of MET25 transcripts in HeLa cells transfected with the pJL-MET construct suggested either poor transcriptional activation or possible transcript instability in this case.

To eliminate the variables introduced by nature of the transient transfection experiment, and to absolutely maximise expression of the MET25 gene, tranfections were employed to produce permanent cell lines which would be expressing the MET25 gene product constitutively, and possibly at higher levels. The construct pLT-MET was used here. Two cell lines expressing low, yet assayable quantities of the MET25 gene product (by Western transfer analysis) were isolated. The lack of sulphydrylase activity measured in protein extracts from these cells finally negated the use of the MET25 gene as a transgene; the behaviour of the MET25 gene product under these conditions was considered to be a sufficiently accurate indicator of its performance in vivo.

3.5.5 Factors contributing to lack of enzyme activity.

Many factors could possibly contribute to the lack of activity of the protein produced in the RRL and cell culture systems. These include (a) production of insufficient protein to allow detection of activity, (b) a requirement for a yeast-specific cofactor for enzyme activity, (c) insufficient or inappropriate post-translational processing of monomers and (d) instability of the mRNA or the protein product. In addition to the haem-related inhibition described in the case of enzyme synthesized in the RRL, one or a combination of these other factors may prohibit the formation of enzymically active tetramer molecules in vitro. These factors may be linked if correct post-translational processing in a yeast environment is required for monomer stability and subsequent tetramer formation. Considering that Yamagata has only assayed OAS sulphydrylase activity in a protein species purified to near homogeneity from a wild-type yeast extract (i.e. the bifunctional OAS-OAH sulphydrylase) and that there have been no other reports of assay of this enzyme activity from protein produced outside the yeast environment, this possibility cannot be ruled out. The factors mentioned in (a) - (d) are discussed separately below.

(a) It is possible to estimate the amount of OAS sulphydrylase enzyme protein produced in the RRL and cell culture systems in terms of the unit of activity of this enzyme. Yamagata has reported that in a crude extract made from a wild-type strain of S. cerevisiae (similar to the strain GP001 used in this study), there is approximately 35 milliunits of activity / mg of protein (Yamagata, 1987; one unit of enzyme activity is defined as the amount of enzyme required to convert 1 μmole of OAS to cysteine per minute). The specific activity of the GP001 extract is less than this (measured to be 8.23 milliunits / mg of protein). Irrespective of the reasons for this, synthesis of cysteine from OAS has even been observed when five-fold less of the GP001 protein extract was used (i.e. when only 2.0 μg of protein from this extract was used in a standard 20 μl sulphydrylase assay, then being at a specific activity of only 1.65 milliunits / mg of protein).

Related to this, and as stated in section 3.4, the results of the Western transfer analyses performed on the RRL and cell culture extracts revealed that the amount of MET25-encoded protein present in 5.0 μg of these extracts was approximately equivalent to that in 2.0 μg of the GP001 extract. However, although the OAS sulphydrylase present in 2.0 μg of the GP001 extract supports synthesis of cysteine from OAS in a standard 20 μl assay, that present in 5.0 μg of RRL and cell culture extracts (ostensibly an equivalent amount of OAS sulphydrylase) is incapable of the same. Evidently, enough protein is present in these extracts to form active tetramers.

If this is so, the lack of sulphydrylase activity measured there could only mean that for activity, the enzyme requires extra processing or a factor which is not present in the mammalian intracellular environment, either during synthesis of the monomers or which otherwise enables tetramer formation and activity. It must be remembered that when DNA from this same cloned MET25 gene is transformed into a yeast mutant lacking OAS-OAH sulphydrylase activity and expressed in the yeast intracellular environment, the wild-type phenotype is restored.

(b) The purity of Yamagata's enzyme preparation suggests that the likelihood of a requirement for any other yeast factor for enzyme activity is low, unless some other small cofactor molecule was copurified with the OAS-OAH

sulphydrylase. Indeed, restoration of enzyme activity (30 %) by simple dialysis of urea-denatured protein monomers (in a Tris-pyridoxal 5-phosphate buffer and in the presence of the reducing agent DTT; see comments below) gives evidence that no small cofactor other than pyridoxal 5-phosphate (found in all cells) is required for enzyme activity.

(c) Interestingly however, there is evidence which supports the possible requirement for post-translational processing of the protein for enzyme activity. First, amino acid analysis of the purified protein revealed the presence of some sugar residues associated with the protein (Yamagata, 1976) suggesting that the protein may be glycosylated. (In contrast, alkaline phosphatase treatment had no affect on enzyme activity, hence phosphorylation is unlikely to be required for activity; S. Yamagata, pers. comm.) The second factor supporting a requirement for post-translational processing is that while the N-terminal residue of the purified protein was determined to be histidine (Yamagata, 1976) methionine is the first codon of the MET25 open reading frame. As histidine residues are found near the N-terminus of the deduced protein sequence at codon positions 4, 11 and 22, it is quite possible that a portion of the N-terminus must be cleaved from the monomer subunit protein in order to form enzymically active tetramers. If this is the case, then incorrect post-translational processing of the monomer subunit is a likely cause of inactivity of the OAS-OAH sulphydrylase produced in vitro. (It should be noted that the monomer and dimer subunits of the OAS-OAH sulphydrylase have no enzyme activity; S. Yamagata, pers. comm.) Other requirements for activity include maintenance of the single cysteine residue (buried in the protein structure of each monomer; not involved in disulphide linkages) in its reduced form for correct folding of the monomer subunit, and the presence of the coenzyme pyridoxal 5-phosphate (found associated with a lysine residue in each monomer).

(d) The possible existence of a sequence within the MET25 gene coding region which is deleterious to its expression in a mammalian environment, either by way of interference of transciption or causing transcript instability, cannot be discounted, especially considering the reported half-life of 10 +/- 1 minute for MET25 transcripts in vivo in the yeast (Sangsoda et al., 1985). Likewise, the resulting protein structure may be unstable when synthesized in a mammalian

environment, especially if the post-translational processing is inadequate. These factors could account for the undetectable level of expression of this gene product in the transient cell culture experiments, and possibly also the low level of expression seen in the two permanent cell lines. (It should be noted that of eight other G418-resistant permanent cell lines isolated following transfection of CHO cells with plasmid construct pLT-MET, none were expressing assayable quantities of the MET25 gene product. This seems a high proportion of "poor-expressers" especially considering the known strength of the RSVLTR as a promoter of transcription, and that the G418-selection conditions favoured cell clones expressing neomycin resistance (and presumably genes associated with this at the integration locus) at a higher level. Although levels of MET25 transcripts in pLT-MET.2 and pLT-MET.9 were not analysed (levels were clearly sufficient to support synthesis of assayable quantities of the OAS-OAH sulphydrylase) this would be consistent with transcript instability, and agree with the result of the primer extension analysis of RNA from pJL-MET transfected HeLa cells, in which no MET25 transcripts were detected.

3.5.6 An alternative to the MET25 gene.

In the absence of sulphydrylase activity in samples of the MET25 gene product synthesized in vitro, successful expression of the cysM gene product of S. typhimurium in mammalian cell culture (A.V. Sivaprasad, manuscript in preparation) has allowed pursuit of the project aim to continue. The cysM gene encodes a protein of 28,000 M_r which is a monomer subunit of the dimeric OAS sulphydrylase-B from this bacterium. Similar to MET25, the gene was identified by complementation analysis using appropriate mutant strains and the complete nucleotide sequence of the coding and 5' and 3' flanking regions determined (A.V. Sivaprasad, unpublished data). In a similar fashion to MET25, the cysM gene product was synthesized in the RRL translation system and assayed for sulphydrylase activity. Like the MET25 gene product however, assay of active enzyme was not achieved. While not tested directly, this was assumed to be for the same reason that the MET25 gene product was not active i.e. interference of the action of the enzyme cofactor, pyridoxal phosphate, by haem in the lysate. Contrary to its yeast counterpart, however, this protein was expressed and

assayed to be enzymically active (in the conversion of OAS to cysteine) in the protein extracts made from transient (COS-1) and G418-resistant permanent (CHO) cell cultures expressing the cysM gene. This is a significant result, considering that the K_m of this enzyme for the substrate OAS is 7.0 mM, marginally poorer than the K_m of the MET25 gene product for OAS, 5.1 mM. This work was done using a eukaryotic expression plasmid similar to pLT-MET, in which the cysM gene was flanked by the RSVLTR promoter and Xenopus ß-globin 5'-untranslated region at its 5' end, and the hGH polyadenylation and other processing sequences at its 3' end. Further details related to this particular plasmid construct, pCS942 (see figure 5.1) will be given in Chapter 5 where the experiments involving coexpression of the cysE and cysM genes are discussed.

Chapter 4

Expression of the cysE Gene.

Chapter 4 : Expression of the cysE Gene.

4.1 Introduction :

In the first of the two reactions of the cysteine biosynthesis pathway described for microorganisms, acetylation of L-serine to form O-acetyl-L-serine (hereafter referred to as OAS) requires the action of an enzyme known as serine acetyltransferase or SAT. In the bacterium Salmonella typhimurium, this enzyme is reported to be a tetramer of identical subunits with a combined molecular weight of 160,000 (Kredich, 1987). In the cell cytoplasm, this usually exists in a M_r 309,000 enzyme complex known as cysteine synthetase, which includes one SAT tetramer and four M_r 36,000 subunits of the enzyme OAS sulphydrylase-A (the OAS sulphydrylase-A subunits exist as two M_r 68,000 molecules in the complex). Interestingly, in a mutant which overproduces the SAT enzyme (Hulanicka and Kredich, 1976) and in another which cannot produce OAS sulphydrylase-A (Hulanicka et al., 1974) serine acetyltransferase activity was measured in the absence of formation of the cysteine synthetase enzyme complex, implying that the SAT enzyme tetramer can function as a separate entity. This is significant for the current work, where in the first instance, the intention is to produce active SAT enzyme in mammalian cells in the absence of OAS sulphydrylase-A.

A study of auxotrophic mutants, designated to be non-functional in cysteine synthesis locus E or cysE, defined a lesion in the first step of the biosynthesis of cysteine from serine (see figure 1.4). By transformation into the cysteine requiring mutant strain JM70, a plasmid clone pRSM8, prepared from an S. typhimurium genomic library and containing a 2.3 kb BamHI/HindIII insert (see figure 3.2 a.) complemented the cysE mutation and hence contained an intact, functional cysE gene (Monroe and Kredich, 1988). Further complementation studies using this full-length clone and various deletion mutants clearly defined a 1.5 kb BamHI/EcoRI fragment as containing the entire cysE gene (E. Kuczek, manuscript in preparation). The complete nucleotide sequence of the cysE gene and 5' and 3' flanking regions was determined (E. Kuczek, unpublished data) and confirmed by comparison with the sequence published for the cysE gene of E. coli. (Denk and Bock, 1987). According to the nucleotide sequence determined

for the cysE gene the enzyme monomer is a polypeptide of 273 amino acids. When translated in an E. coli. transcription/translation system the in vitro protein product of the cysE gene had an apparent molecular weight of approximately 34,000 (E. Kuczek, unpublished data) consistent with previous measurements of the molecular weight of monomer subunits of the SAT enzyme (Kredich et al., 1969).

Following this initial characterisation, the cysE gene was tested with respect to its suitablilty as a transgene for use in the proposed sheep transgenesis progamme. In a manner similar to the MET25 gene, this involved (a) testing the fidelity of translation of the cysE gene coding region in a mammalian environment (namely the RRL translation system) and (b) determining the level of expression of the gene in various mammalian cell lines when the coding region was flanked by appropriate eukaryotic processing sequences. Implicit in each of these was the need to produce an enzymically active gene product capable of the conversion of L-serine to OAS, the first synthesis step in the production of cysteine de novo. Consequently, the ability to detect SAT enzyme activity in small quantities of crude protein extracts was critical. The assay method adapted for use in this is described in Chapter 2. While some of the experimental work described in section 4.2 has been reproduced by the author (especially where quantitation was required) and some new work initiated to further characterize cysE gene expression (i.e in permanently transfected CHO cells), many experiments in expression of the cysE gene in vitro in the RRL and in transient transfection of cell cultures were original work performed by Dr. E.S. Kuczek and are hereby acknowledged. As a result of this, and because much of the experimental detail and rationale of similar experiments have been discussed in the case of the MET25 gene, the presentation of results here will be limited. It should be noted that all theoretical considerations arising from the in vitro experiments described in section 4.2, especially where quantitation is concerned, are original to this work.

Further to the in vitro expression studies (section 4.2) the suitability of the cysE gene for in vivo use was tested through transgenesis in mice. This experimental work is described in section 4.3.

Results.

4.2   Expression of the cysE gene in vitro.

4.2.1   Translation of cysE gene transcripts in the RRL.

A study of expression of the cysE gene in a mammalian environment was initiated by translation of the cysE gene product in the RRL system. To produce SP6 transcripts from the cloned cysE gene, a 979 bp DdeI/EcoRI fragment containing the entire cysE gene coding region along with 89 and 68 bp of 5' and 3' flanking sequence respectively was subcloned from pRSM8 into the transcription vector pGEM-2. Construction of this vector, pGEM-cysE, is described in figure 3.2 a. When SP6 transcripts derived from this clone were translated in the RRL system in the presence of ^{35}S-methionine and the proteins fractionated via SDS-PAGE, fluorography of the resulting gel revealed the presence of a single protein species of approximate molecular weight 34,000 (see figure 3.3, lane 4). In this initial experiment, the cysE coding region appeared to be translated efficiently.

Further, in an assay of SAT activity from a similar but non-radioactive translation reaction, RRL extract containing the cysE gene product converted serine to OAS (figure 4.1, lane 5). [N.B. Conversion of O-acetylserine to N-acetylserine (NAS) by a non-enzymatic acyl migration normally occurs in reactions performed at pH 6.5 or greater (see section 5.2). As a result, NAS rather than OAS is seen as the "product" upon TLC fractionation of SAT reaction mixtures. This is the case in the TLC shown in figure 4.1.] Though not quantitated rigorously, it was estimated that approximately 50 % of the input ^{14}C-serine was converted to OAS in a 30 minute assay, in which approximately 5 μg of total protein from this RRL extract was used as the source of SAT enzyme. This will be considered further in Discussion (section 4.4). No SAT activity was detected in the protein extract from a control RRL translation which contained no added synthetic SP6 RNA (figure 4.1, lane 6).

Figure 4.1 : Assay of the cysE gene product (serine acetyltransferase) synthesized in a RRL translation.

Following translation of cysE SP6 transcripts in a non-radioactive translation reaction in the RRL, the protein product was assayed in a SAT reaction in the presence of ^{14}C-serine (as for all SAT assays) and the assay products fractionated in an ascending TLC in the solvent chloroform : ethanol : acetic acid : water (50:32:10:8) along with the ^{14}C-labelled reference compounds L-serine (^{14}C-SER) O-acetyl-L-serine (^{14}C-OAS) and N-acetyl-L-serine (^{14}C-NAS). A negative control translation lacked RNA, and the positive control was the protein extract made from S. typhimurium strain LT2. The origin and direction of migration of the solvent are indicated in the figure.
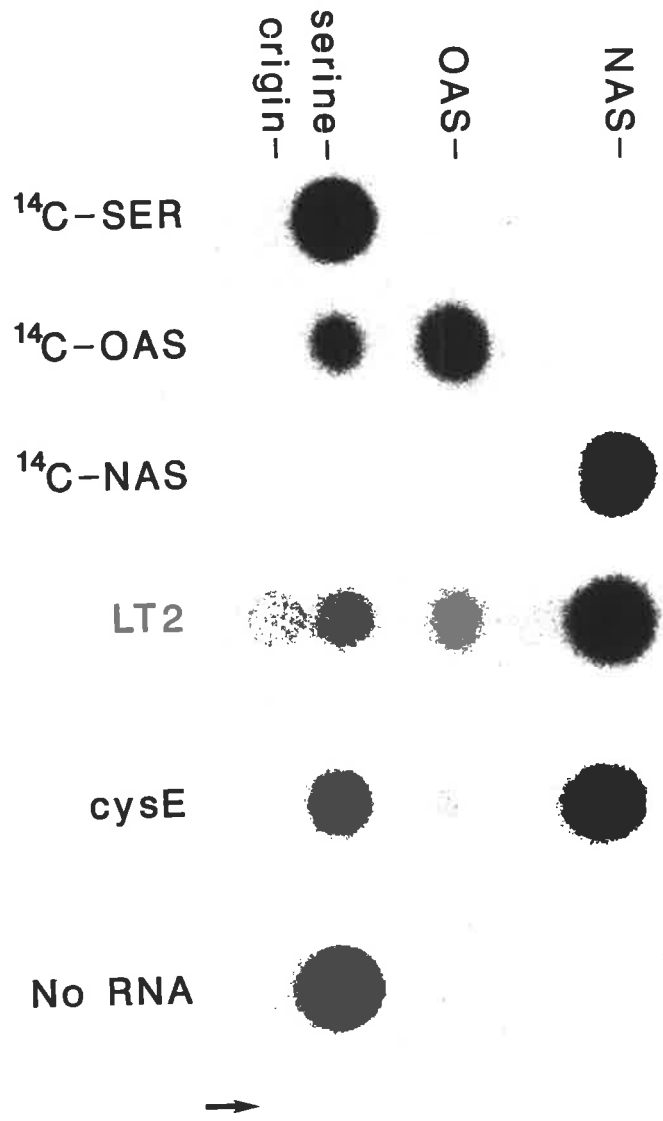
4.2.2 Expression of the cysE gene in mammalian cells.

The successful production of active SAT in the RRL prompted experiments aimed at analysis of expression of the cysE gene in an environment more closely resembling the in vivo situation. Plasmid expression vectors pJL-E and pLT-E (figure 4.2) were constructed in which the cysE gene coding region was flanked by the eukaryotic transcriptional and post-transcriptional processing sequences identical to those found in the corresponding MET25 expression plasmids pJL-MET and pLT-MET (section 3.4 and see figure 3.9). Transient transfection experiments were performed in several mammalian cell lines including HeLa, CHO, COS-1, an SV40-transformed rumen epithelial cell line RESV40 and a primary sheep foetal lung fibroblast cell line CSL503.

Briefly, the cysE gene was judged to be expressed from each plasmid, and in each cell line, by the detection of active SAT enzyme; 10 µg samples of protein extracts made 48 hours after the transfections, were assayed as previously described (figure 4.3). In some assays, all of the input serine was converted to OAS during the 30 minute incubation period, indicating an equivalent, if not higher, specific activity with respect to SAT in these extracts compared to 5 µg of the cysE RRL translation extract. As a rule, the level of SAT activity detected was higher in the protein extracts made from cells transfected with plasmid pLT-E than from those transfected with pJL-E. No SAT activity was detected in the protein extract made from cells of the negative control, where no DNA was included in the transfection.

An additional experiment in which cysE expression was achieved was in the creation of permanent CHO cell lines constitutively expressing the cysE gene product. Following linearization of pLT-E plasmid DNA at the SalI site and transfection of CHO cells by electroporation, six cell lines were isolated after selection in medium containing 350 µg/ml G418 (A.V. Sivaprasad; unpublished data). Detection of high levels of SAT activity in extracts of protein made from subconfluent monolayers of these cells suggested that each of the cell lines was expressing considerable amounts of the cysE gene product. One of these, pLT-E.1, was chosen for further study. (The result of a SAT assay of the protein extract initially made from this cell line is shown in figure 4.3.)

Figure 4.2 (a) : Construction of eukaryotic expression vector pJL-E.

A 979 bp BamHI / EcoRI fragment containing the entire cysE coding region and 89 bp of 5' non-coding and 68 bp of 3' non-coding sequence was isolated from the clone pGEM-cysE (see figure 3.2) and subcloned by blunt-end ligation into the SalI site of pJL4/A+ to create pJL-E. The cloning of this expression plasmid was performed by E. Kuczek.

(b) : Construction of eukaryotic expression vector pLT-E.

To create the expression plasmid pLT-E (i) the same BamHI / EcoRI fragment as in (a) above was subcloned by blunt-end ligation into the unique BamHI site of the base vector pRSVΔH. The essential detail of the 3.1 kb SalI / ScaI transgene fragment isolated from pLT-E for use in the mouse transgenesis experiments described in section 4.3 are shown in (ii). In particular, the position of the restriction enzyme sites in this fragment should be noted. The cloning of this expression plasmid was performed by E. Kuczek.

a.

pJL-E
6.53 Kb

PstI
amp
pBR ori
AvaI/BamHI
EcoRI
BamHI
TGA
A+
cysE
ATG
L
ori
E
BgII
T-antigen

b.

(i)

pLT-E
7.49 Kb

SalI
EcoRI
ATG
SV40
RSVLTR
cysE
TGA
hGH
neo
amp
ScaI
pBR ori

(ii)

SalI
EcoRI   AluI
SmaI   DraI
ScaI
444 b | 336 b | 734 b | 487 b | 1049 b
RSVLTR   cysE   hGH
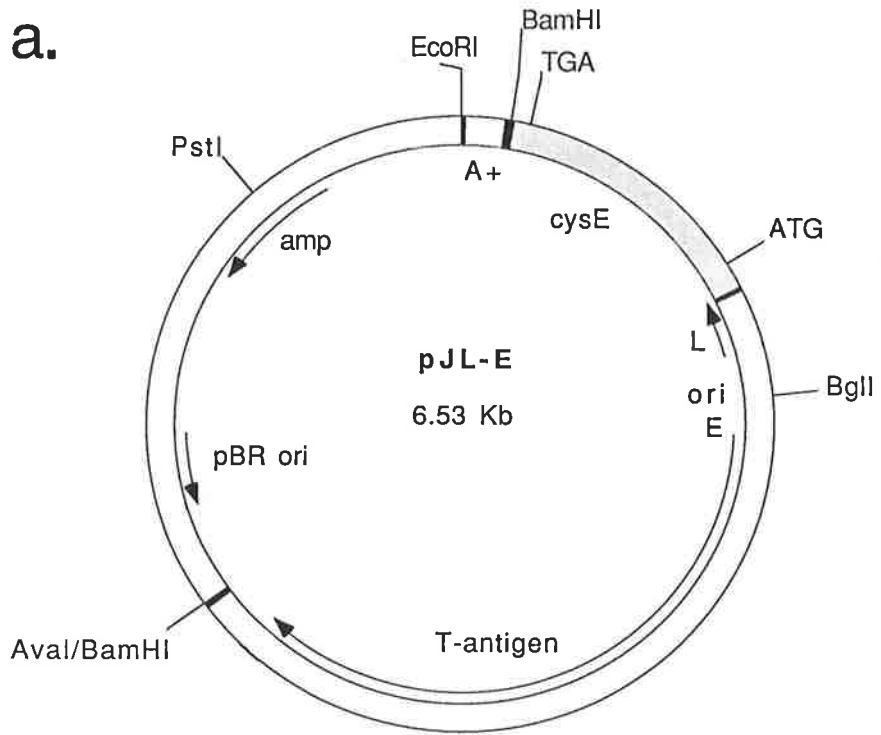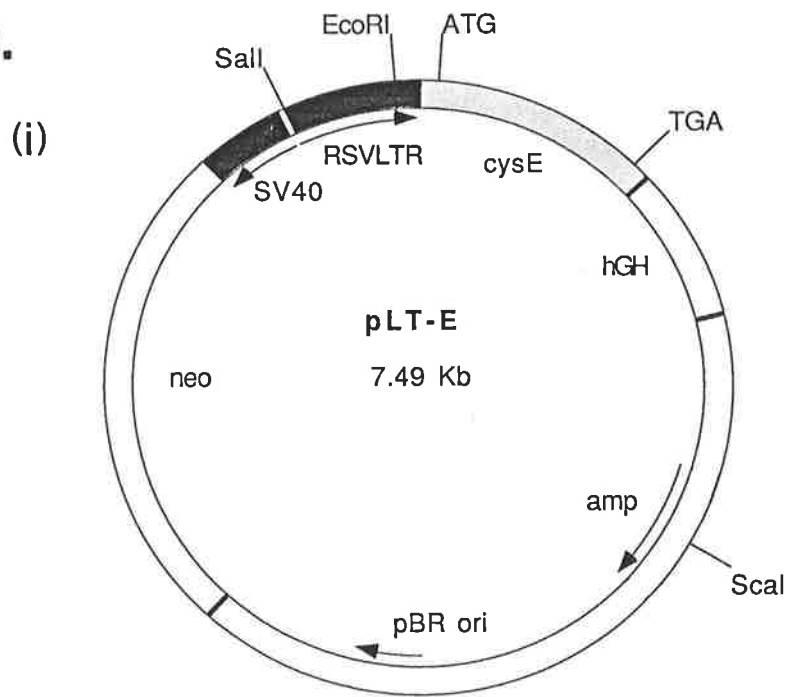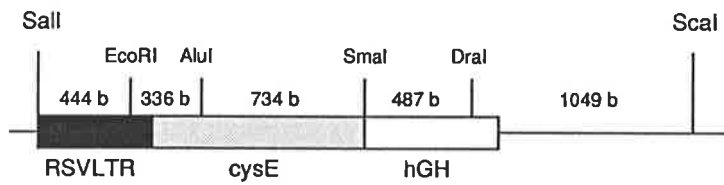
Figure 4.3 : Assay of the cysE gene product expressed in transient and permanent cell cultures.

(a) : The result of the assay of 10 μg protein samples from cell extracts made following the transient transfection of RESV40, HeLa, CHO, CSL503 and COS-1 cells (with constructs pJL-E, pLT-E or No DNA) for SAT activity is shown. The distance of migration of the reference compounds serine (S) O-acetyl-L-serine (O) and N-acetyl-L-serine (N) is given, and the direction of migration of the samples in the ascending TLC fractionation is indicated. Apart from the COS-1 transfection, the other transfections were performed by the persons to follow; E. Kuczek (RESV40, HeLa and CHO transient) R. D'Andrea (CSL503 transient) and A. Sivaprasad (CHO permanent).

(b) : The result of a SAT assay of 10 μg of protein extracts made from the permanent CHO cell lines pRSV.1 (negative control) and pLT-E.1 (stably transfected with the construct pLT-E). The positive control LT2 extract was used in the assay and the reference compounds co-fractionated with the experimental samples in the ascending TLC.

a.

RESV40   HeLa   CHO

NAS-
OAS-
serine-
origin-

S   N   LT2   No DNA   pJL-E   pLT-E   No DNA   pJL-E   pLT-E   No DNA   pJL-E   pLT-E

CSL503   COS-1

NAS-
OAS-
serine-
origin-

S   N   LT2   pLT-E   No DNA

NAS-
OAS-
serine-
cys-4VP-
origin-

S   O   N   GP001   No DNA   pJL-E   pLT-E

b.

CHO

NAS-
OAS-
serine-
origin-

S   O   N   LT2   pRSV.1   pLT-E.1
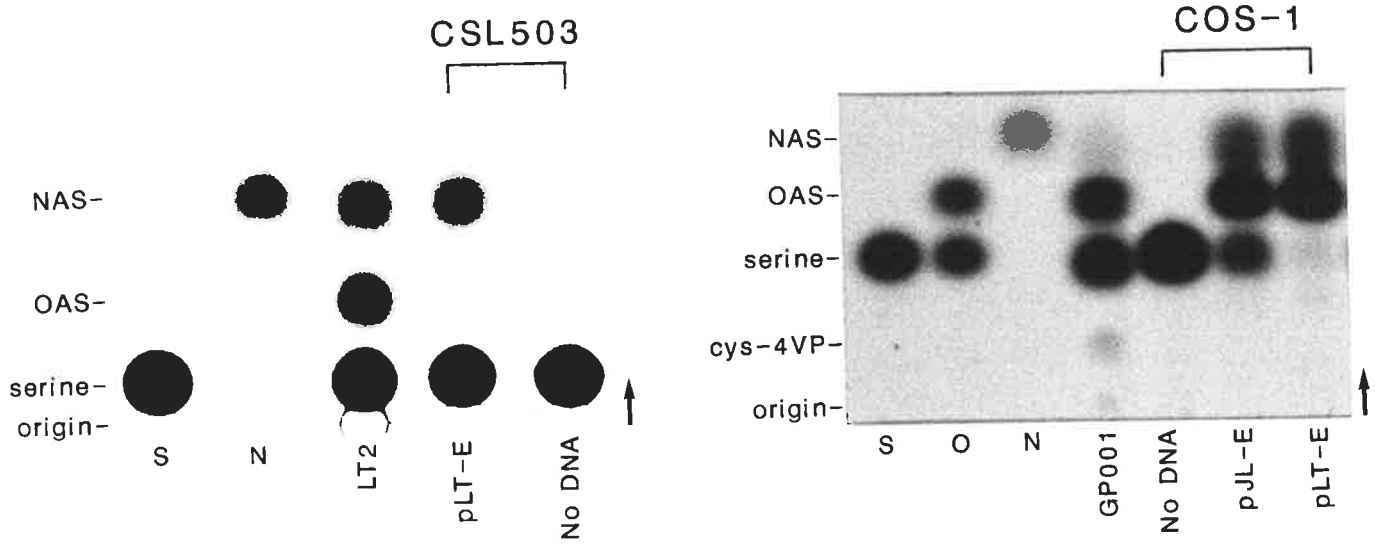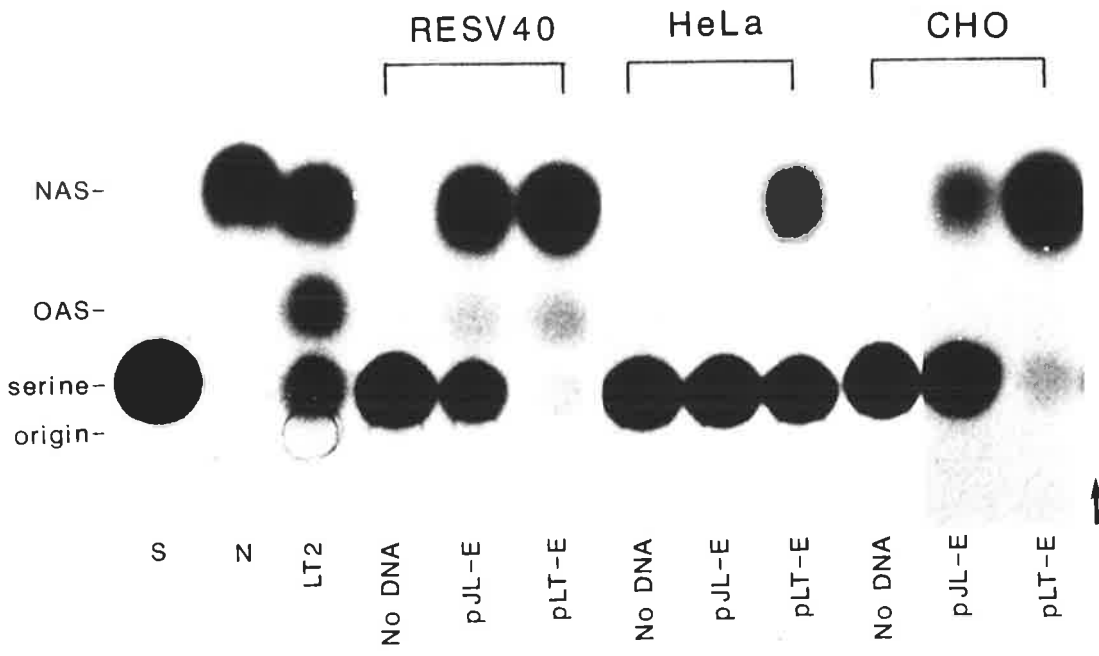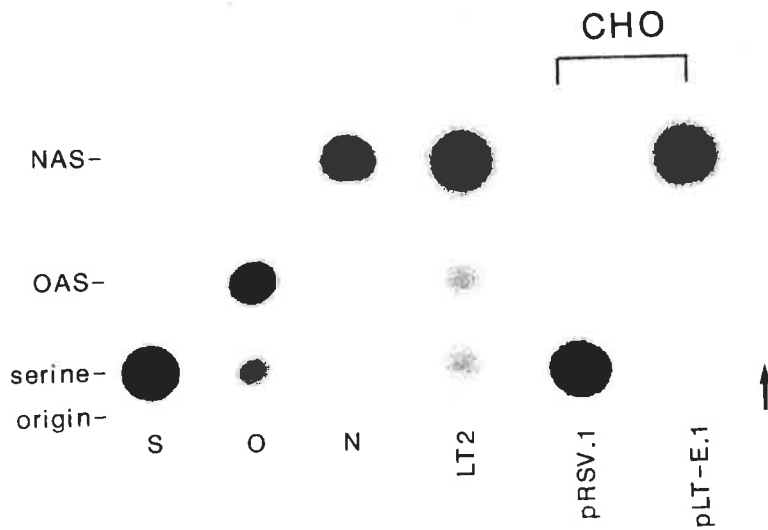
An estimate of the specific activity of the SAT enzyme in the protein extract prepared from pLT-E.1 was made in order to determine the amount of enzyme activity which might be realised in vivo using similar gene promoter and processing sequences. Sub-confluent pLT-E.1 monolayer cultures (approximately 20 million cells) were harvested and a crude whole-cell protein extract (E.1) was prepared. The protein concentration was determined as described in Chapter 2 and a SAT assay time-course reaction was performed over a 60 minute period (figure 4.4 a). Relevant reactants and products resolved in samples in the TLC were excised from the chromatogram and the radioactivity determined by scintillation counting. A graph of % conversion of ^{14}C-serine versus time (figure 4.4 b) revealed the rate of acetylation of serine to be 0.742 nmoles per minute when the rate of OAS synthesis was maximal. Additional calculations related to this rate of catalysis are given in Discussion (section 4.4).

4.3   Expression of the cysE gene in vivo.

Given the successful production of active SAT enzyme when cysE gene constructs were introduced into mammalian cell cultures, the next step was to reproduce this result in vivo via transgenesis. Prior to use of the cysE gene as a transgene in the sheep, where lengthy experiments would be necessary to determine the fidelity of in vivo transgene expression (the gestation period of the sheep is approximately 150 days) it was tested in a more convenient model in vivo system, namely via transgenesis in mice. The successful expression of a gene construct in transgenic mice using the same promoter had previously been reported (Overbeek et al., 1986).

4.3.1   Generation and identification of founder transgenic mice.

In order to generate transgenic mice carrying the RSV/SAT/hGH hybrid gene, a 3.1 kb SalI/ScaI fragment from the plasmid pLT-E was purified as described in Chapter 2 and microinjected into fertilised ova. This DNA fragment contains 535 bp of the RSVLTR promoter region, 979 bp including the cysE gene of S. typhimurium, 622 bp of 3' flanking sequence including the polyadenylation and associated 3'-processing sequences from the human Growth Hormone gene and 905 bp of pUC19 sequence (figure 4.2 b (i) and (ii)). All experimental

Figure 4.4 : SAT assay time course analysis of extract E.1 (from G418-resistant cell line pLT-E.1).

(a) : A SAT assay time course was performed in a 200 μl volume using 10 μg of E.1 protein and including 1 mg/ml BSA (to prevent the enzyme inactivation observed previously upon dilution of the E.1 extract 60-fold; data not shown) 1 mM acetyl-CoA, 10 mM Tris.HCl (pH 6.5; the use of Tris.HCl at this lower pH will be discussed in section 5.2) and 1.03 mM L-serine (a combination of 30 μM ^{14}C-serine and 1.0 mM of non-radioactive serine, giving a final 34.3-fold dilution of the ^{14}C-serine, to enable quantitation of SAT enzyme activity at a point when the reaction rate is constant). Over a 60 minute incubation period, time sample aliquots (10 μl) were taken at 0, 1, 2 and 5 minutes, then every 5 minutes for the remainder of the time. Reactions were stopped by heating at 100 ^{0}C for 3 minutes, protein removed by centrifugation and 5 μl of the supernatant loaded directly onto a TLC sheet. Reactants and products were fractionated in the TLC using the solvent of Checchi (1961) and autoradiography of the dried chromatogram performed at room temperature for 3 days. The reference samples serine (S) OAS (O) and NAS (N) were cofractionated with the time course samples.

(b) : Following autoradiography of the chromatogram produced in (a) densitometric analysis was performed. The SAT activity of the E.1 extract (% conversion) is graphed here as a function of time. * 1.0 % conversion of L-serine to O-acetyl-L-serine (OAS) per minute by 10 μg of the E.1 extract in this SAT assay is equivalent to the conversion of 0.206 nmoles of L-serine per minute. As one unit of SAT activity is defined as the SAT enzyme required to convert 1.0 μmole of L-serine to OAS per minute, then the specific activity of this extract is approximately 74.2 milliunits/mg of protein. (The specific activity of a crude protein extract made from S. typhimurium wild-type strain LT2 is about 30 milliunits/mg.)

a.

NAS—
OAS—
serine—

origin—

S   O   N   0   1   2   5   10 15 20 25 30 35 40 45 50 55 60

|_____| Time (minutes)
 Standards

b.

SAT Assay Time Course : E.1 Extract.
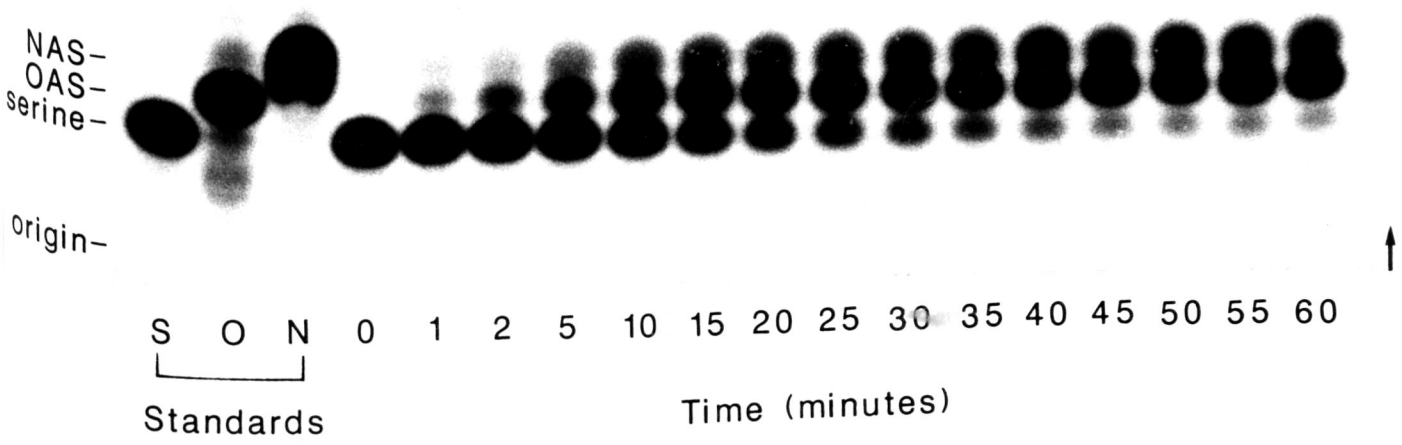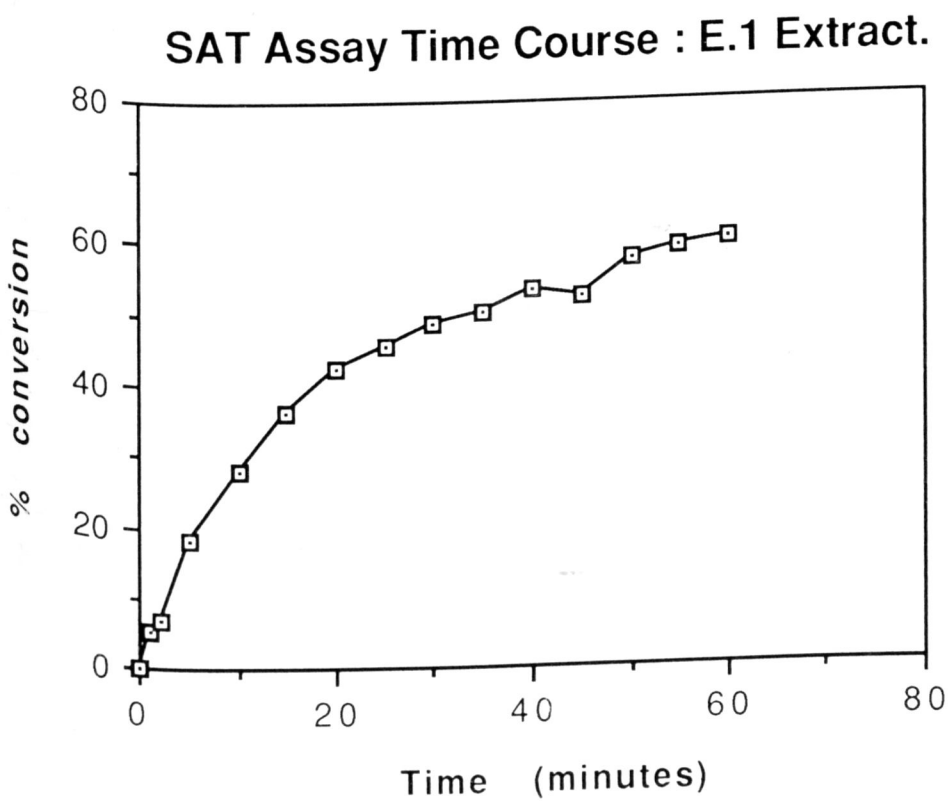
aspects related to the collection, fertilisation and microinjection of ova and the subsequent generation of mouse pups were carried out by P.J. Verma (see acknowledgements) according to the methods of Brinster et al. (1981) and Hogan et al. (1986). At three weeks post-partum, the author began analysis of the pups thus generated.

Nine founder (F_0) transgenic mice, designated E-11, -16, -17, -20, -21, -23, -25, -26 and -27 were detected amongst 30 newborn pups by dot-blot hybridization analysis of genomic DNA prepared from tail tissue (figure 4.5 a). This represented a level of 30 % transgenesis among the newborn mice. Within these 9 F_0 animals, copy numbers of the integrated transgene varied from one to approximately 25 copies per diploid genome (hereafter, copy numbers given refer to the number of copies of the integrated transgene per diploid genome). Analysis of the genomic arrangement of the transgene in the F_0 mice by Southern blot hybridization of EcoRI-digested tail DNA (there is only one EcoRI restriction site in the transgene fragment, contained within the RSVLTR promoter region) indicated that the 3.1 kb fragment had been integrated at random sites in the genome (figure 4.5 b). Founder E-21 appeared to have three separate integration loci. This was indicated by detection of the presence of at least five EcoRI fragments in this Southern and upon analysis of the F_1 offspring (see section 4.3.3). Sites of multiple copy integration contained tandemly arranged head to tail arrays of the fragment (figure 4.5 b). In the case of 4 of the F_0 mice (E-11, -17, -23 and -25) binding of the hybridisation probe to similar sized high molecular weight DNA species suggested incomplete cleavage of the integrated transgene sequences by EcoRI. This was unexpected, but observed again in a repeat Southern transfer using newly prepared DNA which had also been purified by spermine precipitation prior to the enzyme digestion. Similar high molecular weight species were detected in the EcoRI-digested DNA of E-26. No hybridizing species were detected by Southern blot analysis of the EcoR1-digested genomic DNA of F_0 mouse E-16 or in the EcoRI-digested genomic DNA of a non-transgenic control mouse (figure 4.5 b). In the case of E-16, initially presumed to have only a single copy of the transgene in its genome, a sensitivity problem was assumed because in a repeat Southern transfer experiment in which a larger quantity of DNA was used, the transgene sequences were detected (see figure 4.7

Fig. 4.5 (a) : Dot-blot hybridisation analysis of genomic DNA from cysE F_0 transgenic mice.

5 µg of genomic DNA from the F_0 mice generated in the cysE mouse transgenesis programme was transferred to Genescreen membrane and probed with a ^{32}P-labelled Nick-translated 3.1 kb SalI / ScaI DNA fragment isolated from pLT-E (the same fragment as that used in the microinjections). A copy control series made by combination of appropriate amounts of the plasmid pLT-E with 5 µg of genomic DNA prepared from a non-transgenic mouse was transferred to the membrane along with the test DNA samples.Detection of 5 pg of the 3.1 kb transgene fragment diluted into 5 µg of mouse genomic DNA (the mouse genome size is about 3×10^9bp) gives a signal equivalent to that of a single-copy gene sequence of that size. The filter was washed in 0.2 x SSC at 65 ^{0}C and autoradiography carried out at -80 ^{0}C for 12 hours.

(b) : Southern transfer analysis of EcoRI-digested genomic DNA prepared from the tail tissue of cysE F_0 mice transgenic mice.

Approximately 4 µg of EcoRI-digested genomic DNA from the F_0 cysE transgenic mice was analysed by the Southern transfer hybridisation protocol. The DNA was transferred to Zetaprobe membrane and hybridised with the probe described in (a). The filter was washed in 0.2 x SSC at 65 ^{0}C and autoradiographed at -80 ^{0}C for 48 hours. The position of the 3.1 kb unit-length fragment detected in the DNA of E-26, -27 and the copy control lanes is shown (◄) along with that of the high molecular weight DNA species detected in the DNA of E-11, -17, -23, -25 and -26 (◁). The migration distance and size of the HindIII-digested Lambda marker DNA fragments is on the left of the figure.

(c) : SAT assay analysis of tail protein extracts from the cysE F_0 transgenic mice.

10 µg of protein was assayed for SAT activity in the presence of ^{14}C-serine and the reaction products fractionated by ascending TLC in the Checchi solvent. The test assay samples are shown on the left and the control assays of the pLT-E.1 extract (E.1; +) and a protein extract from a non-transgenic mouse (-) are on the right along with the reference compound samples. Autoradiography of the resulting chromatogram was for 3 days at room temperature.

a.

| 0 | 0.5 | 1 | 2 | 5 | 10 | 25 | 50 | 100 |

-Copy no. controls

-Fo mouse DNA

11 16 17 20 21 23 25 26 27

b.

kb

23.1-

9.4-

6.6-

4.4-

3.6-

2.8-

◄3.1 kb

11 16 17 20 21 23 25 26 27   0   1   5   25

Fo mouse DNA
samples

Copy no.
controls

c.

NAS-

OAS-

serine-

origin-

11 16 17 20 21 23 25 26 27   S   O   N   E.1   -

Fo mouse assay samples

Controls
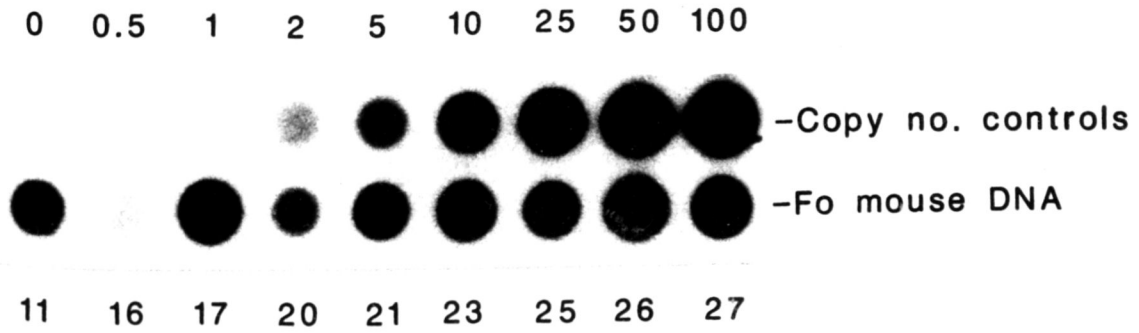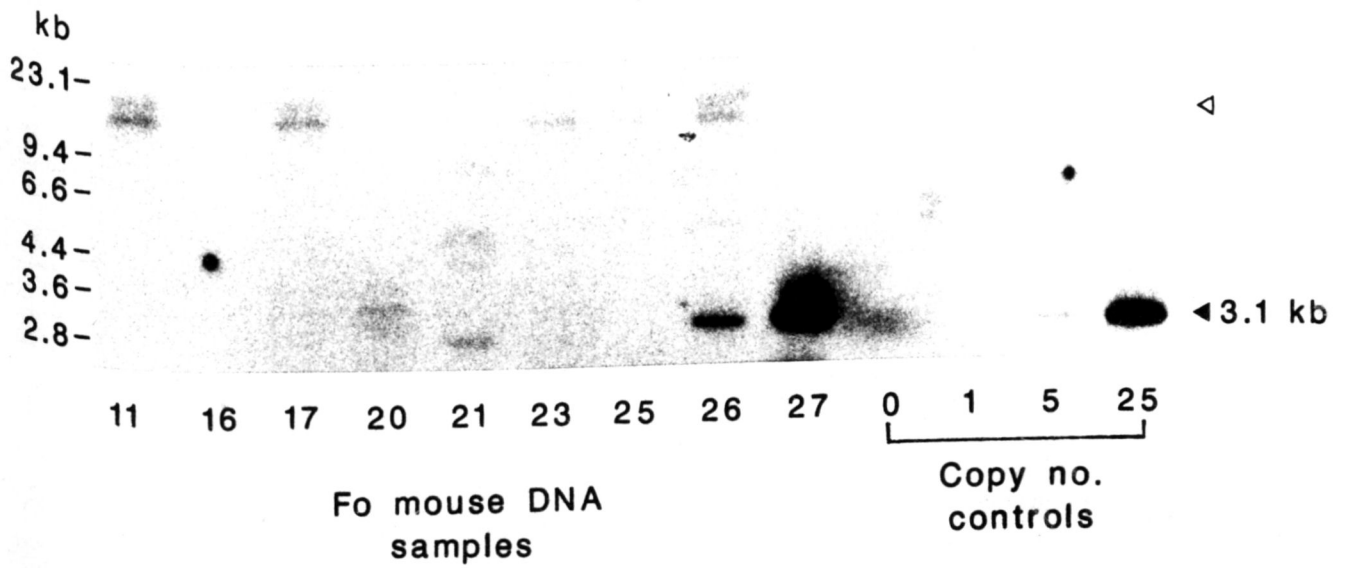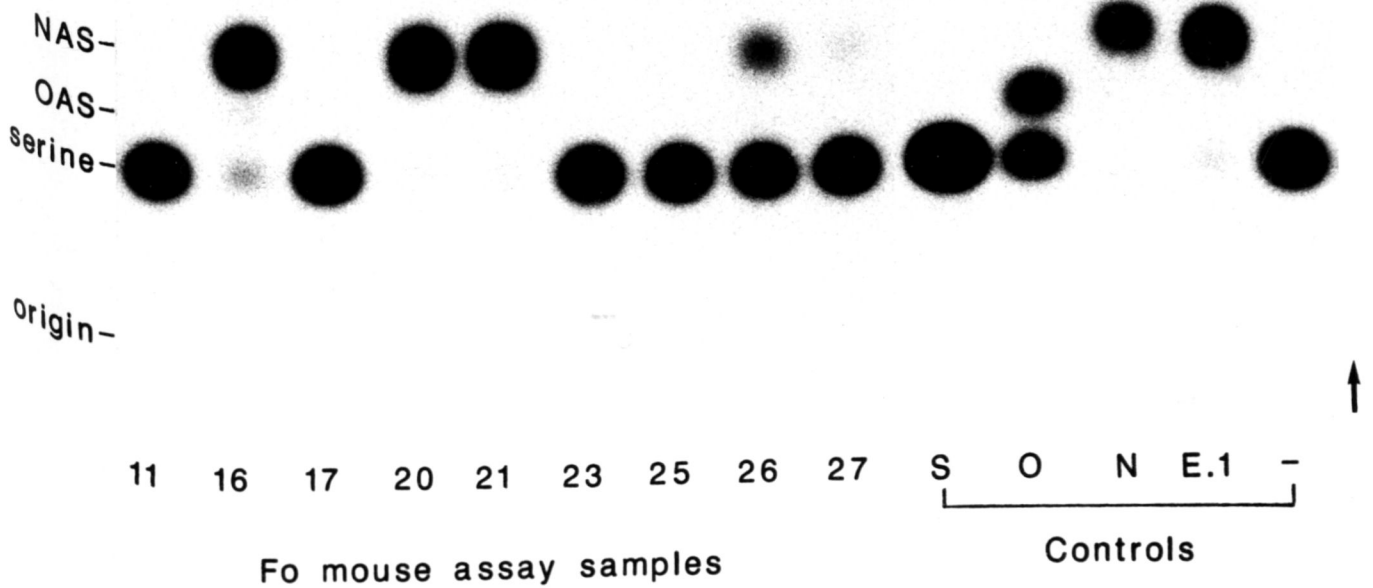
c). Moreover, expression of the transgene was detected by SAT assay of a protein extract made from the tail of this mouse (see figure 4.5 c).

4.3.2 Analysis of transgene expression in F_0 mice.

Protein extracts prepared from the tail tissue of each of the 9 F_0 transgenic mice (according to the modified method described in Chapter 2) were assayed for SAT activity as an initial test for gene expression. This tissue is reported to provide a good representation of mesoderm-derived cell types in which the RSVLTR promoter directs a high level of gene expression (Overbeek et al., 1986). As shown in figure 4.5 c, only 5 of the F_0 mice exhibited detectable levels of SAT enzyme activity in tail tissue; three at very high levels (E-16, -20 and-21) and two at much lower levels (E-26 and -27).

After the establishment of mouse lines derived from the founder males E-16 and E-21 (see section 4.3.3) the level of expression of the transgene in the tissues of 18 selected organs and body structures (hereafter referred to as "tissues") from 8 of the F_0 animals (F_0 mouse E-26 had died prior to this analysis) was tested by assay of the respective protein extracts for SAT activity (figure 4.6). The F_0 animals which exhibited no SAT activity in their tail tissue (E-11, -17, -23 and -25) had no SAT enzyme activity in any other tissue. In the case of three of the other F_0 animals (E-16, -20 and -21) a significant level of SAT activity was detected in 11 of the 18 tissues examined. High levels of SAT activity were observed in the heart, leg muscle, diaphragm, sternum, tail, foot (hind) and skin (dorsal) and lower levels of enzyme activity were also evident in other tissues, notably the lung, brain, thymus and eye. Only trace amounts of SAT enzyme activity were detected in the other tissues (including the stomach and small intestine) excepting the liver, where none was detected. In the case of the remaining F_0 transgenic, E-27, only low levels of SAT activity were detected in some tissues. In this mouse, only the seven tissues listed previously, in which a high level of SAT activity was seen in the other F_0 mice (i.e. the heart, leg muscle etc.) exhibited any SAT activity, and even this was at the limit of detection in this assay. No SAT activity was detected in any of the tissues of a non-transgenic F_0 mouse, E-1.

Figure 4.6 : SAT assay analysis of protein extracts from 18 selected tissues from each of the cysE F_0 transgenic mice.

In these SAT assays, 20 μg of protein prepared from each of the tissues listed, and for each of the F_0 cysE transgenic mice except E-26 (which died prior to this analysis) was used. In each assay, ^{14}C-serine was used as the substrate. The TLC's were each autoradiographed for 3 days at room temperature, excepting for that of E-27, which was a 7 day exposure. The position of the origin of sample loading is shown (see the left side of the figure). The reference samples and the E.1 positive control were co-fractionated with the tissue assay samples of each of the F_0 mice.

origin-

origin-

origin-

origin-

origin-

origin-

origin-

origin-

E-11

E-16

E-17

E-20

E-21

E-23

E-25

E-27

¹⁴C-SER
¹⁴C-OAS
¹⁴C-NAS
E.1
Heart
Spleen
Liver
Lung
Brain
Thymus
Eye
Leg muscle
Kidney
Diaphragm
Pancreas
Stomach
Sternum
Small Intestine
Testis/Ovary
Tail
Foot
Skin

Another feature noted in comparing SAT activity in the tissues of F_0 animals was that an identical pattern of SAT activity was seen amongst these mice, with respect to both tissue type and general level of activity in the various tissues. For example, SAT activity was usually always detected in the leg muscle and tail, and the activity in these two tissues was invariably higher than that seen in other tissues such as the lung or eye.

4.3.3   Heritability of the transgene.

The usefulness of a transgene is restricted by the degree to which it is able to persist in the genome of the host species, be accurately expressed and passed from one generation to the next. To test the heritability of functional cysE transgenes the male founders, E-16 (originally judged to be a single gene copy, high-expressor) and E-21 (a multiple gene copy, multi-locus, high-expressor) were mated to normal females and the F_1 offspring analysed with respect to inheritance of the genotype and phenotype of their F_0 parents. The two resulting mouse lines will now be discussed separately.

In the case of E-16, only 2 from 48 offspring, one male, 16.20, and one female, 16.47 (the rate of transmission of the transgene was only 4.2 %) were determined to be transgenic by dot-blot analysis (figure 4.7 a). The dot blot data revealed that the genomes of these mice contained more copies of the transgene (approximately 2 to 5 copies) than the F_0 parent, E-16 (1 copy). When tail protein from these F_1 mice was initially assayed for SAT activity (figure 4.7 b) each animal was determined to be producing a level of active SAT enzyme higher than that of E-16. Each of these features of the F_1 generation suggest chimaerism in the F_0 mouse E-16. In order to compare the F_0 and F_1 mice further, Southern blot analysis of EcoR1-digested genomic DNA prepared from the tail tissue of E-16, 16.20 and 16.47 was performed (figure 4.7 c). The presence of three hybridizing species in the digested DNA of F_0 mouse E-16 suggests that more than a single copy of the 3.1 kb transgene fragment is contained in its genome and that the arrangement of these is in a head to tail array. In fact, judging from the band intensities, it may be expected that there are 2 copies of the transgene fragment in the genome of F_0 mouse E-16; the fainter band at 7.0 kb and the 2.7 kb band could represent genomic flanking sequences bearing the 5' (444 bp RSVLTR)

Figure 4.7 (a) : Dot-blot hybridization analysis of genomic DNA from F_1 and F_2 offspring derived from the cysE transgenic founder E-16.

5 µg of genomic DNA from the F_0, F_1 and F_2 mice (line E-16) was transferred to Genescreen membrane and hybridised with the 3.1 kb SalI / ScaI probe isolated from pLT-E. The filter was washed at 0.2 x SSC / 65 ^{0}C
then autoradiography performed for 8 hours at -80 ^{0}C. A copy control series was included in the analysis and the position of these and the test mouse DNA samples is shown in the table immediately beneath the autoradiogram.

(b) : SAT assay analysis of tail protein extracts from the F_1 and F_2 transgenic offspring of E-16.

Protein extracts (10 µg) from F_0 mouse E-16, F_1 mice 16.20 and 16.47, and from F_2 mice 2047.2, 2047.10 and 2047.12 were assayed for SAT activity in the presence of ^{14}C-serine and the reaction products fractionated by ascending TLC. The assay of E-16 was performed in duplicate. The reference compounds serine (S) OAS (O) and NAS (N) were cofractionated in the TLC and their distance of migration is shown on the left side of the figure. The TLC was autoradiographed for 3 days at room temperature.

(c) : Southern transfer hybridisation analysis of EcoRI-digested genomic DNA from F_0 mouse E-16 and F_1 mice 16.20 and 16.47.

10 µg of EcoRI-digested genomic DNA from E-16, 16.20 and 16.47 was transferred to Zetaprobe membrane and hybridised with the 3.1 kb SalI / ScaI transgene fragment probe. EcoRI-digested DNA from a non-transgenic mouse was used as the negative control (-). The filter was washed at 0.2 x SSC / 65 ^{0}C and autoradiography carried out at -80 ^{0}C for 9 hours. The position and size of major fragments detected in this analysis (◄) and that of the marker fragments in a mixture of HindIII digested lambda DNA and EcoRI-digested SPP-1 DNA are indicated in the figure.

(d) : Use of a cysE-specific probe in hybridisation to the E-16 F_0 / F_1 mouse genomic Southern filter.

A 1.0 kb EcoRI / SalI cysE-specific fragment isolated from pGEM-cysE (see figure 3.2) was ^{32}P-radiolabelled by Nick-translation and used to reprobe the same Southern filter as was used in (c). The filter was washed and autoradiographed as described in (c). The position and size of the major hybridisation products detected is indicated on the right side of the figure.

a.
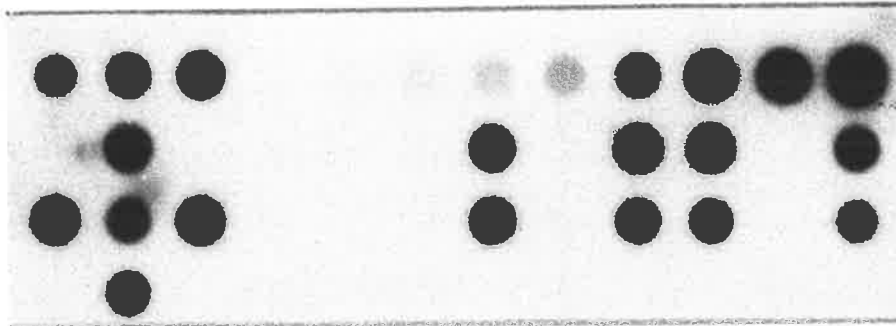
E-16	16.20	16.47	-	0	0.5	1	2	5	10	25	50	_Copy no. controls
.1	.2	.3	.4	.5	.6	.7	.8	.9	.10	.11	.12	
.13	.14	.15	.16	.17	.18	.19	.20	.21	.22	.23	.24	
.25	.26	-	-	-	-	-	-	-	-	-	-	

F₂ mouse DNA

b.

NAS-
OAS-
serine-

origin-

S   O   N   E-16 E-16 16.20 16.47 2047.2 2047.10 2047.12

F₀   F₁   F₂

c.

kb
23.1-
9.4-
6.6-   ◄ 7.0 kb
4.4-

2.3-   ◄ 3.1 kb
              ◄ 2.7 kb
1.5-   ◄ 1.6 kb

- E-16 16.20 16.47

d.

- E-16 16.20 16.47

◄ 7.0 kb
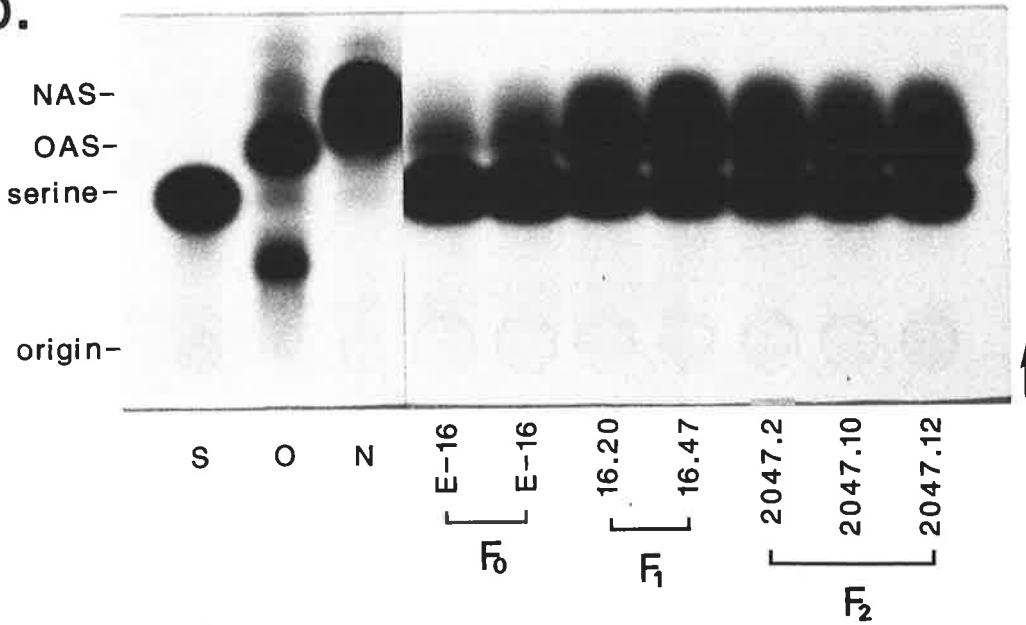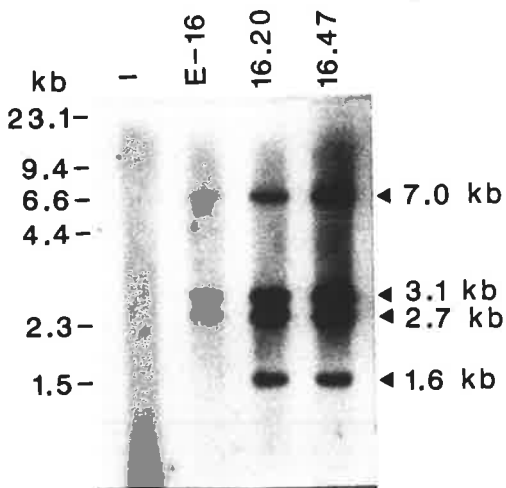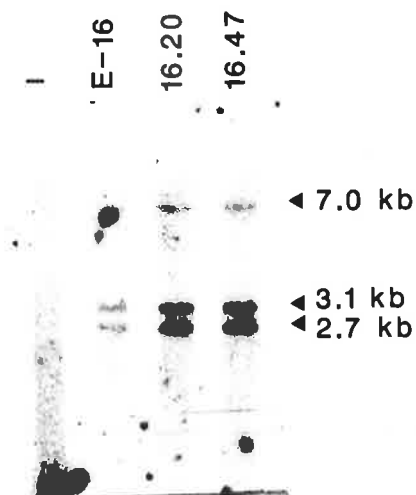
◄ 3.1 kb
◄ 2.7 kb

and 3' (2606 bp cysE/hGH) portions of the transgene fragment, whilst the band at 3.1 kb is a unit length of the transgene (expected in an EcoRI digest of a multi-copy, head to tail array of this fragment). When a longer autoradiographic exposure of the Southern filter failed to reveal any other faint bands which could represent junction fragments (containing genomic DNA flanking a transgene locus) the possibility of the existence of two separate loci was ruled out.

If the genomic arrangements of the 3.1 kb transgene fragment in the F_1 offspring are examined, they appear to be identical to each other and the same as E-16, except for an additional hybridizing species approximately 1.6 kb in size. When this Southern blot was reprobed with a 1.0 kb EcoRI/SalI fragment (isolated from pGEM-cysE; see figure 3.2 a) containing only the cysE gene sequences, all fragments were detected except the 1.6 kb fragment (figure 4.7 d). This was also the case when the blot was reprobed with a 487 bp SmaI/DraI fragment (isolated from pLT-E; see figure 4.2 b (ii)) containing only the hGH 3'-end sequence (data not shown). These results suggest that the 7.0 kb fragment may contain the small 444 bp RSVLTR portion and also at least one whole copy of the 3.1 kb transgene fragment. (Taking into account the possibility of a lower efficiency of transfer of this larger fragment in the Southern, the intensity of the band in the autoradiogram suggests that probably only one whole transgene copy is present within the 7.0 kb band.) Hence the genome of the chimaeric F_0 mouse E-16 in fact contains at least 3 copies of the transgene (though obviously not in all cells). If this is the case, the 7.0 kb fragment may contain one uncleavable EcoRI site, which was inherited as such by the F_1 offspring. Secondly, the 2.7 kb fragment, detected by all three probes, obviously represents transgene DNA 3' to the EcoRI site, including the cysE gene and hGH sequences. In addition, it is clear that the 1.6 kb fragment must be derived from either the RSVLTR 5'-end and/or plasmid vector sequences. Though unusual, this extra band gives further evidence of the chimaeric nature of the founder E-16. Unless due to a sequence-specific gene rearrangement during spermatogenesis (see section 4.4) where further such rearrangements could be expected in the F_2 generation, the presence of this extra sequence should be revealed by a similar Southern analysis of EcoRI-digested genomic DNA prepared from the testes of E-16.

To examine the heritability of the transgene further, the F_1 mice 16.20 and 16.47 were mated to each other to produce an F_2 generation with respect to the founder, E-16. As could be expected, many of the offspring (13 from 26; a transmission rate of 50 %) from this mating were transgenic (dot-blot hybridisation analysis; figure 4.7 a). In addition, when tail protein from each of these mice was assayed, all of the 13 transgenics expressed a high level of SAT activity in this tissue. This suggested that the functional transgene had remained intact in the F_2 animals. A crude determination of the SAT activity present in the tail tissue of the F_2 mice 2047.2, 2047.10 and 2047.12 is shown in figure 4.7 b for comparison with those of E-16, 16.20 and 16.47. In a separate SAT assay time course experiment, the specific activities of the tail protein extracts of F_0 mouse E-16, F_1 mice 16.20 and 16.47 and F_2 mouse 2047.2 were determined using 20 µg of protein. The results of this analysis are graphed and tabulated in figure 4.8. Expressed as milliunits of SAT activity per milligram of protein, the values for E-16, 16.20, 16.47 and 2047.2 were 7.6, 31.7, 62.4 and 35.4 milliunits / mg respectively. Indeed, a higher specific activity of SAT (4 to 8-fold) is seen in the protein extracts of the F_1 and F_2 mice when compared with that of the F_0 mouse E-16. Also, the specific activity of SAT in the tail protein extract of the F_2 mouse 2047.2 is of the order of that of its F_1 parents 16.20 and 16.47, and is higher than the SAT activity measured for E-16 tail protein extract. Another feature of this analysis obvious from the graph is the decreasing reaction rate observed in all of the protein samples assayed in this fashion. Considering especially the case of the assay of E-16 tail protein, where the enzyme activity is intrinsically low (and falling substrate concentrations are thus less likely to limit the rate of reaction) it appears that the enzyme becomes unstable in an extended assay performed under these conditions. Whilst this interferes somewhat with the determination of initial reaction velocities (and causes slight underestimation of the specific activities of the tail protein extracts with respect to SAT activity) it is assumed to be an artifact of the assay system, most likely due to incomplete inhibition of proteases during preparation of the tail protein extract. This is a highly probable cause, as this phenomenon was always more noticeble in the case of extracts initially containing less of the cysE gene product.

Figure 4.8 : SAT assay time course analyses of protein from transgenic mice E-16, 16.20, 16.47 and 2047.2.

(a) : 20 μg of protein from tail protein extracts of the E-16 F_0, F_1 and F_2 mice was analysed for SAT activity in time course reactions, in the presence of 30 μM ^{14}C-serine and 1.0 mM non-radioactive serine. Samples were taken every 10 minutes over a 30 minute period. Following TLC fractionation of the products and autoradiography of the chromatograms for 3 days, densitometric analysis of the resulting autoradiograms was carried out. The conversion of serine to OAS is plotted here as a function of time. * 1 % conversion per minute is equivalent to the conversion of 0.206 nmoles of L-serine to OAS per minute, as in the SAT time course analysis of the CHO cell pLT-E.1 extract (see figure 4.4).

(b) : The optimum SAT activity measured in the time course assays described in (a) is expressed in the table in milliunits of SAT activity per milligram of protein, where one unit of activity is defined as the amount of enzyme required to convert 1.0 μmole of L-serine to OAS per minute.

a.

SAT Assay Time Course : Mouse Tail Protein Extracts.
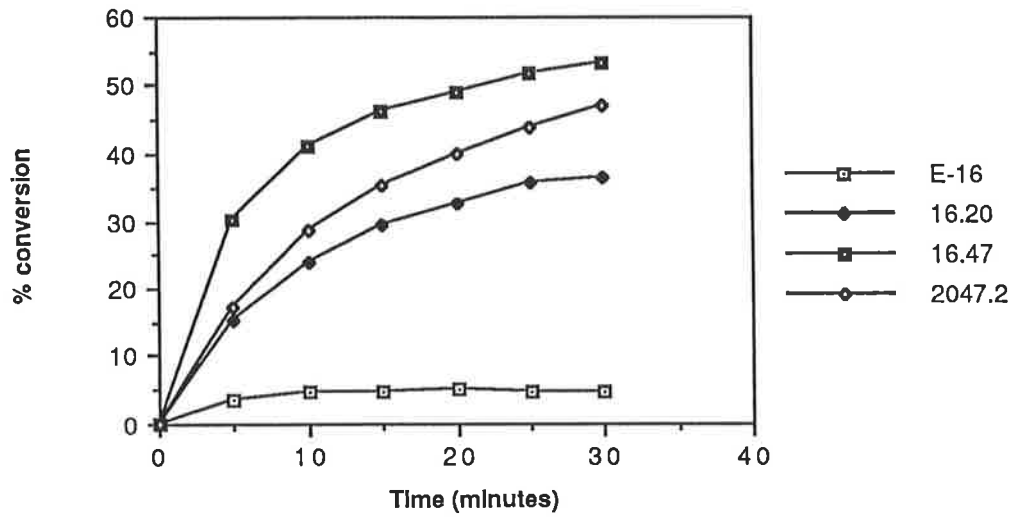
b.

Mouse	SAT activity (milliunits/mg)
F_0 E-16	7.6
F_1 16.20	31.7
F_1 16.47	62.4
F_2 2047.2	35.4

A similar analysis of the male founder E-21 was carried out by mating this with normal females. In the F_1 generation, a very high number (40 of the 57 mice; the rate of transmission of the transgene in this case is 70.2 %) were determined to be transgenic (figure 4.9 a). Transgene copy numbers were widely variant, and again different to that of the founder animal. Analysis of all of the F_1 transgenic mice by SAT assay of tail protein revealed only 20 of these to be expressing the transgene and a number of different levels of SAT activity (four main expression phenotypes including the non-expressing phenotype; see figure 4.9 b) were evident. This is discussed further in section 4.4.4. An F_2 generation of these mice was not produced.

To conclude this study of heritability of the cysE hybrid transgene, the F_1 mouse 16.47 (derived from F_0 mouse, E-16) and the F_1 mouse 21.4 (derived from F_0 mouse, E-21) were sacrificed and expression of the cysE gene in 18 tissues examined by SAT assay and compared to that seen in the tissues of the founder. The result of this analysis for mouse 21.4 is shown in figure 4.10. Comparing the SAT activity detected in the tissues of the founder and the tissues of the respective offspring, an identical pattern of activity of the enzyme is seen, with respect to both tissue type and general level of activity in the tissues. The assay result from the SAT analysis of the tissues of the F_0 transgenic non-expresser, E-23, is shown alongside this. Additionally, in the tissues of the F_1 mouse 16.47, there is a higher level of SAT enzyme activity compared to those of its chimaeric parent, E-16 (data not shown). Clearly, the tissue-specific pattern of enzyme activity originally observed was preserved in subsequent generations.

4.3.4 Quantitation of SAT enzyme activity in the tissues of transgenic mice.

The results of assays for SAT activity in protein extracts made from the tissues of transgenic mice revealed different levels of expression of the cysE gene. The pattern and level of expression seen in the tissues examined was similar to that reported previously by Overbeek et al. (1986). To more accurately compare the current findings with those reported earlier and to determine the maximum level of SAT activity which might be expected and achievable in vivo, the protein extracts made from four tissues (leg muscle, heart, brain and liver) of

Figure 4.9 (a) : Dot-blot hybridization analysis of genomic DNA from F_1 offspring derived from the transgenic founder E-21.

5 µg of genomic DNA from F_0 mouse E-21 and each of the 57 F_1 offspring was transferred to Genescreen membrane and hybridised with the 3.1 kb SalI / ScaI transgene fragment. The filter was subsequently washed at 0.2 x SSC / 65 ^{0}C and autoradiography carried out at -80 ^{0}C for 16 hours. A copy control series was included in the analysis and the position of these samples with respect to F_0 and F_1 mouse DNA samples is indicated in the figure.

(b) : SAT assay analysis of tail protein extracts from E-21 F_1 offspring 21.8 - 21.15.

20 µg of tail protein was assayed for SAT activity in a 60 minute time course reaction performed as previously described. The products were fractionated by ascending TLC alongside reference compounds serine (S) and NAS (N), and a positive control reaction (a cysteine synthesis reaction in this case, in which all SAT assay components were used along with the SAT and OAS sulphydrylase activities present in the E.1 and M.1 CHO cell extracts respectively; see section 5.2 and figure 5.1). The chromatogram was autoradiographed for 3 days at room temperature. Mice designated as genotype (-) are non-transgenic while those of genotype A, B, C or B/C are transgenic mice with different expression phenotypes (resulting from different genotypes created by segregation of the different transgene loci present in the F_0 mouse E-21).

a.

0   0.5   1   2   5   10   25   50   100   F₀ E-21

Copy no. controls

F₁ E-21 mouse DNA samples

b.

−   C   −   B/C   −   B   C   A   −Expression phenotype

NAS−
OAS−
serine−

cysteine-4VP−
origin−

S   N   E.1+M.1   21.8   21.9   21.10   21.11   21.12   21.13   21.14   21.15

F₁ E-21 assay samples
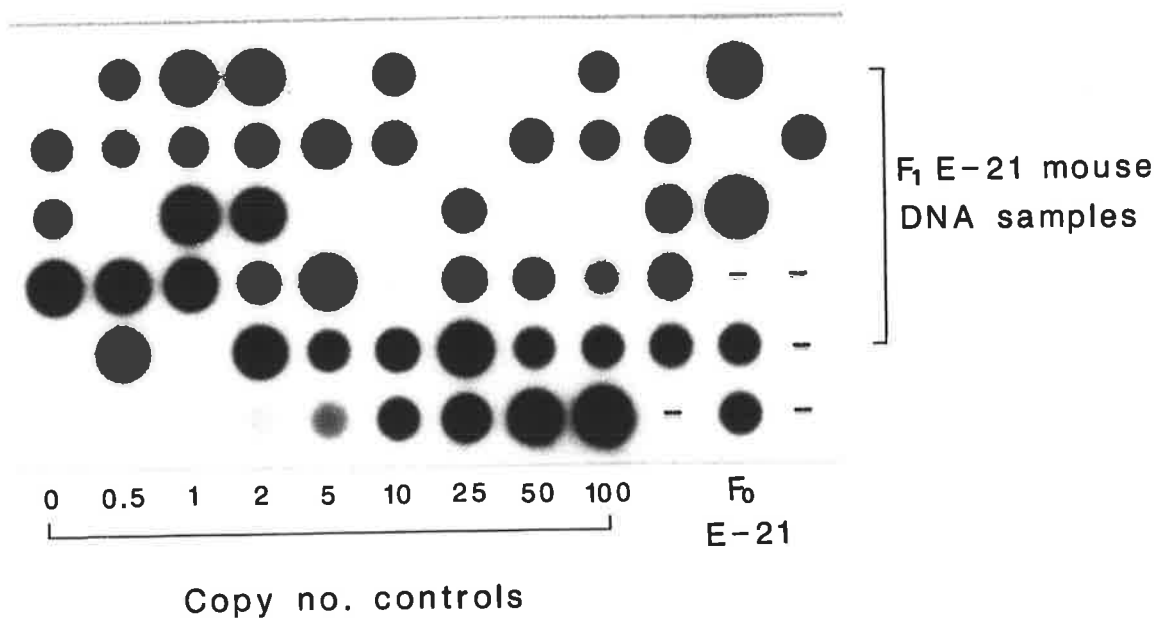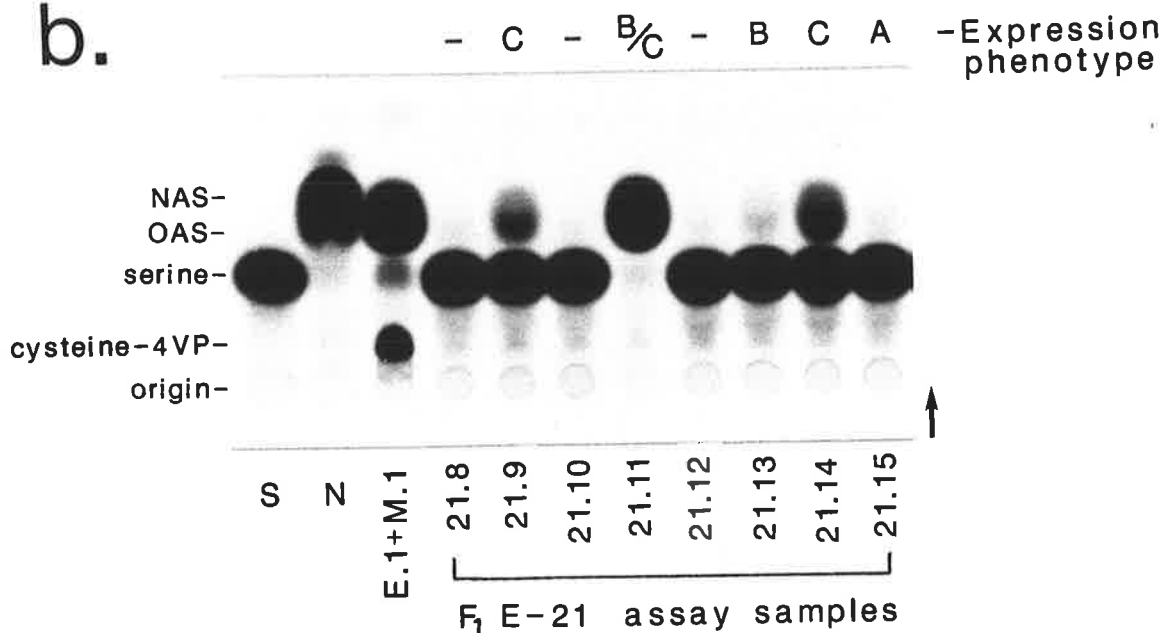
Figure 4.10 : Comparison of SAT activity in protein extracts from tissues of the cysE transgenic F₀ mouse E-21 and the F₁ offspring 21.4

20 μg of protein from each of 18 selected tissues of transgenic mice E-21, 21.4 and an F_0 cysE transgenic (non-expressing) mouse E-23 was analysed by SAT assay and the reaction products fractionated by ascending TLC alongside the reference compounds serine (^{14}C-SER) OAS (^{14}C-OAS) and NAS (^{14}C-NAS), and a positive control (E.1 extract SAT assay). The origins of sample loading and direction of migration of the solvent (Checchi) are shown in the figure. Autoradiography was carried out for 3 days at room temperature.

a.

^{14}C-SER | ^{14}C-OAS | ^{14}C-NAS | + | Heart | Spleen | Liver | Lung | Brain | Thymus | Eye | Leg Muscle | Kidney | Diaphragm | Pancreas | Stomach | Sternum | Small Intestine | Testis | Tail | Foot | Skin

F_o E-21

origin-

b.

F_1 E-21.4

origin-

c.
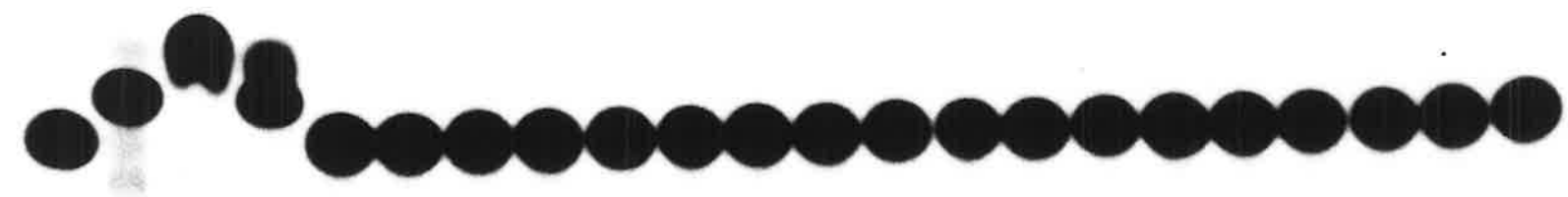
F_o E-23

the transgenic F_1 male, 21.4, were assayed in separate time course reactions. Calculations of maximum enzyme activities were made from the amount of OAS detected in assay time samples taken from reactions performed under conditions where neither of the substrates, acetyl-CoA (1 mM) or L-serine (1.03 mM, including 30 μM ^{14}C-serine) was limiting. The results of the time course reactions for the leg muscle and heart are graphed (figure 4.11 a) and the maximum enzyme activity calculated for each tissue is tabulated (figure 4.11 b). This analysis revealed maximum SAT activity levels of 386, 70.9 and 0.024 milliunits/mg of protein in the tissues of the leg muscle, heart and brain respectively. No SAT activity was detected in the liver protein extract.

Discussion.

4.4   The suitability of cysE as a transgene.

When the results of the in vitro and in vivo analyses are considered the suitability of the cysE gene from S. typhimurium for use as a transgene in the sheep is unquestionable. What follows is a discussion of the experimental observations which prove this, including some calculations related to the activity of the SAT enzyme produced in these systems.

4.4.1   Activity of the cysE gene product in a mammalian environment.

In the first instance, the cysE gene product is translated well in a mammalian system, the RRL, without any modification to the coding or flanking DNA sequences. Whilst this translation efficiency may be improved (e.g. as in the case of the MET25 gene; see section 3.3) it enables the production of enough SAT protein monomers to provide active enzyme tetramers. In assays of the activity of such an RRL extract, it is estimated that approximately 50 % (i.e. 300 pmoles) of the input ^{14}C-serine is converted to OAS in a 30 minute assay, in which 5 μg of total protein from a cysE RRL translation extract is used as the SAT enzyme source. Where one unit of enzyme is defined as that amount able to convert one μmole of serine to OAS per minute (Kredich and Tomkins, 1966) there is the equivalent of 10 μunits of SAT enzyme in 5 μg of protein from this RRL translation extract (i.e. the specific activity of this protein extract with

Figure 4.11 : SAT assay time course analysis of protein from the leg muscle, heart, brain and liver of F_1 transgenic mouse 21.4.

(a) : 20 μg of protein from the leg muscle, heart, brain and liver tissue extracts of cysE F_1 transgenic 21.4 was assayed for SAT activity in a 60 minute time course reaction in the presence of both ^{14}C-serine and non-radioactive serine as previously described. Time samples were withdrawn from the reactions at 0, 1, 2,and 5 minutes then after every 5 minutes until the 60 minute time point. Following fractionation by ascending TLC in the Checchi solvent and autoradiography for 3 days at room temperature, the autoradiogram was subjected to densitometric analysis. The calculated % conversion of L-serine to OAS in only the leg muscle and heart samples is graphed as a function of time, and 1 % conversion of serine to OAS per minute is again equivalent to conversion of 0.206 nmoles of L-serine per minute.

(b) : The optimum SAT activity measured in the time course assays of all tissues of E-21.4 described in (a) is expressed in the table in milliunits of SAT activity per milligram of protein, where one unit of activity is defined as the amount of enzyme required to convert 1.0 μmole of L-serine to OAS per minute.

a.

SAT Assay Time Course : Heart and Leg Muscle Protein from E-21.4
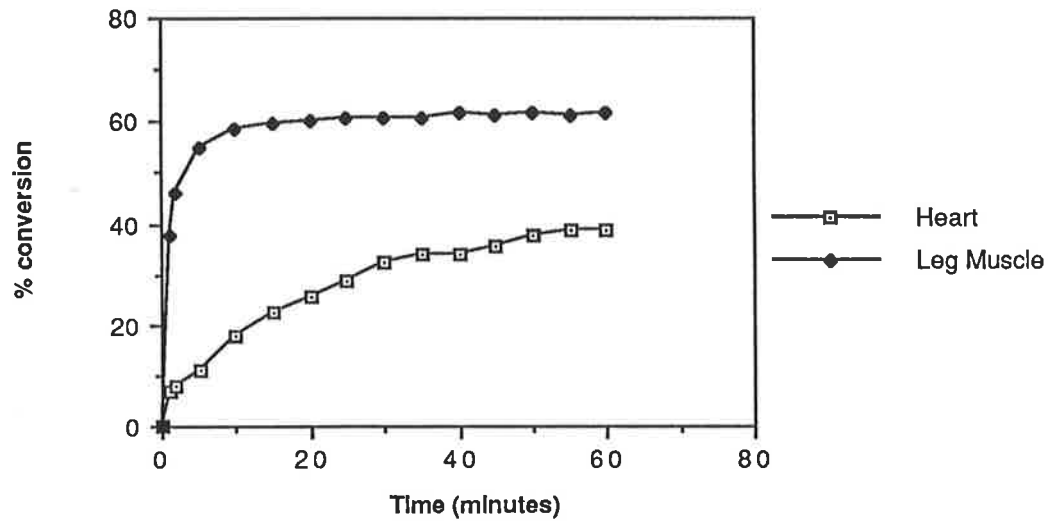
b.

Mouse tissue (E-21.4)	SAT activity (milliunits/mg)
Leg muscle	386
Heart	70.9
Brain	0.024
Liver	-

respect to SAT activity is 2.0 milliunits/mg of protein). According to the figure of 26 units/mg of purified SAT enzyme (Kredich and Tomkins, 1966) 5 μg of protein from such a non-radioactive translation theoretically contains 385 pg of SAT enzyme monomers. In the 20 μl SAT reaction, this obviously provides a concentration high enough to form the tetramers required for enzyme activity. (The concentration of monomers - of approximate native molecular weight 27,300 according to a 273 amino acid polypeptide - in the 20 μl SAT reaction is of the order of 0.71 nM, and so tetramers are at about 0.18 nM in this assay).

Considering that this assay for SAT activity may not be optimal (though quite sensitive and performed at a higher temperature than that used in the earlier work of Kredich et al.; 1966, 1969) the calculation made above may be an underestimation of the true mass of SAT in the RRL, hence also of the monomer concentration. Bearing this in mind, the production of active SAT enzyme in vitro in the RRL supported two suppositions related to activity of the SAT enzyme. First, the monomer subunits of SAT are indeed identical (and encoded by a single gene, the cysE gene) and secondly, SAT subunits can form a tetramer which is enzymically active in the absence of the OAS sulphydrylase-A enzyme with which they are usually associated in the cysteine synthetase complex. The evidence which suggests that tetramers, or possibly at least dimers, are formed by association of molecules of the cysE gene product is indirect. There are only two active sites in the 160,000 M_r tetramer (Kredich et al., 1969) indicating that a polymeric structure must be formed for the active sites to be created.

Another encouraging aspect of the production of active SAT in this mammalian in vitro system, is that if this prokaryotic protein has any specific post-translational processing requirements, these can and have been met by the resident eukaryotic processing mechanisms (c.f. the case of the MET25 gene product). This was demonstrated more convincingly by the successful expression of active cysE gene product in transient cell culture transfections in which many different cell types were transfected using the plasmids pJL-E and pLT-E, where the cysE gene is under the control of the SV40 late and RSVLTR promoters respectively. The presence of higher levels of SAT activity in protein extracts made from cultures transfected with the pLT-E plasmid suggested that the RSVLTR promoter exhibited greater transcriptional activity than the SV40 late

promoter in this system. Also, the synthesis of active SAT in the sheep primary cell line gave a positive indication of the likely success of expression of the cysE gene as a transgene in the sheep.

Expression of the cysE gene in a manner more closely resembling the in vivo situation was achieved through creation of permanently transfected CHO cell lines constitutively expressing the cysE gene. When assayed, 10 μg of a protein extract (E.1) made from one of these cell lines (pLT-E.1) was able to convert 0.74 nmoles of serine to OAS per minute. Compared to the unit definition, this implies that in 10 μg of E.1 protein, there are 0.74 milliunits of active SAT enzyme (this represents a mass of approximately 1.5 ng of enzyme molecules) and that the specific activity of this crude extract is 74 milliunits/mg of protein. Interestingly, the specific activity of a crude SAT extract made from Salmonella typhimurium is a little less than half this, at 30 milliunits/mg of protein (Kredich and Tomkins, 1966).

In terms of cells, it is estimated that of the original number harvested to make the E.1 extract, the protein from about only 100,000 cells was included in the assay. If the SAT activity present in this number of cells was complemented by a non-limiting level of OAS sulphydrylase activity (as is the case in vivo in S. typhimurium; Kredich, 1987) then 0.74 nmoles (92 ng) of cysteine would be synthesized per minute. Thus to synthesize a minimum of 1.5 g of cysteine per day (required for significant improvement to wool growth; refer to section 1.3) 1.13×10^9 cells would be required. An estimation of the rumen surface area required to provide this number of cells (assuming the surface of one rumen epithelial cell to be 200 square μm, and the presence of four layers of actively synthesizing cells in this epithelium) is approximately 565 square cm. Considering the size and highly convoluted surface of this organ in the sheep, availability of cells is not considered a problem. This demonstrates that the level of SAT activity required to produce a minimum of 1.5 g of cysteine per day (in 1.13×10^9 epithelial cells lining the rumen) could theoretically be achieved, through expression of the cysE gene in a manner as seen in the permanent cell culture system, where the gene is integrated into the host cell genome and efficiently expressed with the aid of a powerful eukaryotic promoter.

4.4.2 Expression of the cysE gene in transgenic mice.

The ultimate proof of the suitability of cysE as a transgene was provided by the expression of enzymically active cysE gene product in a tissue-specific manner in transgenic mice. Generation of many normal founder animals, containing various numbers of copies of the transgene fragment indicated that integration of these particular DNA sequences caused no aberration of the normal pattern of development. While integration of these foreign sequences has caused no difficulty, it is also evident that expression of the cysE transgenes caused no gross morpholigical aberration of the normal phenotype.

When tested by SAT assay for transgene expression, 5 of the 9 F_0 mice were shown to produce active SAT enzyme. The level of SAT activity detected in the tail tissue of these mice was indicative of expression levels in other tissues of the animal. Analysis of SAT activity in 18 tissues revealed a tissue-specific pattern of cysE gene expression almost identical to that reported in mice made transgenic with an RSVLTR CAT gene construct (Overbeek et al., 1986). This was also true with respect to the general level of expression seen in particular tissues. In tissue from the heart, leg muscle, diaphragm, sternum, tail, foot and skin, high levels of SAT activity were measured. In addition, there was a lower level of SAT activity detectable in many other tissues. The selective, yet consistent tissue differences in gene expression can be explained by considering the relative preponderance of mesodermally-derived cell types within the tissues, and that the preferential site of expression of the RSVLTR promoter has been reported to be in cells of mesodermal origin (Overbeek et al., 1986). Hence it is also possible that trace amounts of gene expression are detected in a diverse range of tissues because connective tissue, which is associated with many tissues, is of mesodermal origin. This would account for the low levels of SAT enzyme activity detected in non-mesodermal tissues such as the lung and small intestine.

Preservation of the general pattern of tissue specificity of the RSVLTR promoter is also supported by the results of analysis of levels of SAT activity in the heart, leg muscle, brain and liver of the F_1 mouse 21.4 in section 4.3.4. If quantitation of SAT activity in these tissues is an accurate reflection of the level of gene expression (this is assumed from the findings of Overbeek et al., 1986) then near identical transcriptional constraints apply to expression of the cysE gene in

vivo as apply to expression of the CAT gene in the earlier work. This favours the suggestion that the information targeting the expression of the integrated transgene to particular tissues is indeed wholly contained within the RSVLTR promoter fragment, and more importantly, that the cysE coding region does not have any deleterious effect upon interactions between the RSVLTR promoter and the transcriptional machinery.

4.4.3   The effects of the host genome on expression of the transgene.

Transgenic founders not expressing detectable levels of the cysE gene product in their tail tissue were also found not to be expressing the transgene in any other tissue. Interestingly, the integrated transgene sequences of the four non-expressing mice, and some of the copies integrated into the genome of E-26, appeared to be resistant to cleavage by the restriction enzyme EcoRI, for although the genome of each of these mice contained multiple copies of the transgene sequence, the unit length 3.1 kb transgene fragment was not resectable. This was also the true in the case of one of the integrated copies in the transgene locus in mouse line E-16, and of a similarly constructed RSVLTR linked-gene construct when integrated into the genome in various permanent CHO cell-lines (see section 5.4.2). Considering that one EcoRI cleavage site exists in the transgene fragment, both the lack of expression of the transgene in these adult mice and resistance to digestion by EcoRI would be consistent with methylation of the DNA at the site of integration. EcoRI cannot cleave a recognition sequence which is methylated at the second A residue (i.e. GAmATTC). Previously, it has been reported that in vitro methylation of an EcoRI* site (AATT) situated 79 bp from the cap site of a cloned herpes simplex virus thymidine kinase gene prior to microinjection of the construct into the nuclei of thymidine kinase deficient cells abolishes expression of this gene (Waechter and Baserga, 1982). However, no EcoRI site-specific methylase enzymes have been described in eukaryotes, hence this explanation seems implausible.

Despite this, methylation at other sites could account for the lack of expression in these individuals. In particular, gene expression governed by LTR promoters is known to be succeptible to methylation at CpG dinucleotides, such

as those found in HaeII, HhaI and HpaII sites that are situated either within, or in flanking DNA upstream or downstream of the promoter (Feenstra et al., 1986; Mietz and Kuff, 1990). Interestingly, two such HhaI sites, spaced 80 bp apart are found 5' to the U3 region of the RSVLTR promoter (which contains the TATA box and enhancer sequences), the closest of these being about 141 bp upstream. Otherwise, a chromosomal position effect, such as that described by Al-Shawi et al. (1990) may be responsible for inactivation of the transgene.

In regard to the transgene structure, other factors including sequence-directed DNA rearrangements, random mutations or damage due to breakage and rejoining during integration may have led to abolition of the EcoRI sites in each transgene copy integrated into the genomes of the multi-copy F_0 transgenics E-11, -17, -23 and -25. Depending upon the sequence requirements of the RSVLTR promoter, loss of the EcoRI site (situated only 20 bp away from the TATA box and about 50 bp from the start-site for transcription) as a result of any of the means described, could either seriously impair or abolish transcription. With reference to the latter of these possibilities, it is conceivable that the structure of a transgene integration intermediate, containing two or more of the RSVLTR regions in tandem, may lend itself to some form of sequence-directed rearrangement, such as homologous recombination between adjacent LTR regions (Varmus and Swanstrom, 1985; Fujiwara and Mizuuchi, 1988). Given that the creation of 1-LTR and 2-LTR circles has been shown to occur in the nucleus (Brown et al., 1987) this or a similar mechanism may account for the appearence of the unexpected 1.6 kb DNA species in the genomes of mice of the E-16 line. Although the possibility of site-specific integration of the RSVLTR promoter sequence into preferred target sites exists (Shih et al., 1988) the role of this in relation to integration of the transgene sequences in use here remains to be investigated. Nonetheless, the use of LTR sequences in future transgenesis experiments will be discussed further in Chapters 5 and 6.

Considering the limited amount of data gathered in this study regarding the levels of gene expression and gene copy numbers detected in the founder animals and their offspring, it is not possible to determine whether expression of the cysE hybrid transgene exhibits copy-number dependence. However, as might be expected, even in examining the expression levels and copy numbers of only

the male founders E-16 (chimaera) E-20 (single locus, 4 copies) E-21 (triple locus chimaera) and E-27 (single locus, 20 copies) it is clear that expression of the cysE transgene is integration-site dependent. The three integration loci of E-21, which display different levels of cysE gene activity, are a good example of this.

4.4.4 Heritability of functional transgenes.

Heritability of the transgene was demonstrated by analysis of the offspring of F_0 mice E-16 and E-21. Whilst the structure of the integrated transgenes in the F_2 mice remains to be analysed, the function of the transgene was preserved in the F_1 and F_2 generation mice derived from E-16. In addition, copy numbers higher in the offspring than in the F_0 individuals suggest that the latter were each chimaeric animals, made transgenic by integration of the transgene fragment into genomic DNA after the single cell stage of development of the mouse embryo. (Evidently, many such chimaeric animals have been produced in the transgenesis programmes currently operating here.) The only transgenic F_1 offspring of E-16, 16.20 and 16.47, displayed a high level of SAT activity in their tail tissue and subsequent F_2 transgenic mice had a similarly high level of SAT activity.

Analysis of heritability of the transgene in the F_0 mouse E-21 was only pursued to the level of the F_1 generation, due to the possibility of three separate integration loci in the genome of this animal. In the F_1 offspring generated, a high rate of transmission of the transgene (70.2 %) was to be expected considering the likely segregation of these three loci during spermatogenesis. The different levels of expression of the transgene evident in the F_1 animals (judged by SAT assay to be of four main expression phenotypes, designated types A, B, C and B/C) indicated that one of the loci (A) was transcriptionally inactive (e.g. as in F_1 mouse E-21.15; see figure 4.9 b) and that the other two loci (B and C) displayed low and moderate levels of activity respectively (e.g. as in F_1 mice E-21.13 and E-21.14 respectively; see figure 4.9 b).

The additive effect of the expression of the cysE gene at the two active loci, as seen in E-21 and in many of the F_1 offspring containing at least these two loci (e.g. as in F_1 mouse E-21.11, designated an expression type B/C mouse in

figure 4.9 b) was significant. In fact, compared to the low (approximately 1-5 %; locus B) or moderate (approximately 10-20 %; locus C) proportions of conversion of serine to OAS, the near 100 % catalysis seen in the SAT assays of mice with a combination of each of these loci (as in the F_0 parent, E-21) was disproportionate. This suggests that above a threshold concentration of enzyme monomers, a greater cooperativity of association, and hence catalytic capacity, exists. If this is the case, one might expect, for example, that even a transgenic strain homozygous for the low-level expression locus B, might exhibit a level of SAT activity higher than that expected by simple addition of the gene activities from each separate B locus. This highlights the fact that expression of the cysteine biosynthesis genes, which each encode only monomer subunits of a particular enzyme, must be maintained at a level sufficient to satisfy the requirements for the physical association of these monomers in the formation of catalytically active species.

The expression of active SAT enzyme in vivo in the transgenic mice discussed is of primary importance in one other respect, that being the maximum level of enzyme activity achievable in this environment. Although a similar quantitation was entertained in the case of the permanent CHO cell line, pLT-E.1, a study of the performance of the cysE gene in vivo was considered more valuable to the long-term objective of this thesis. With this in mind, four tissues from the F_1 transgenic mouse 21.4 were analysed. According to the maximum level of SAT activity measured in the protein extract made from the leg muscle of this mouse (386 milliunits/mg of protein) it may be concluded that in 1.0 mg of protein from the epithelial cells lining the rumen of the sheep, the SAT enzyme from a similarly well-expressed transgene could support the synthesis of 386 nmoles of OAS per minute (i.e. 0.56 mmoles per day). In relation to cysteine synthesis, this means that only 22.3 mg of protein of this specific activity would be required to enable the synthesis of 1.5 g (12.4 mmoles) of cysteine per day. Even at a far more moderate specific activity of SAT enzyme, such as 10 milliunits/mg of protein, only about 1.0 g of total cell protein from the rumen epithelium would be required, a relatively small proportion of the protein of the rumen epithelium considering the size of this organ in the sheep.

In summary, the production of active SAT enzyme by expression of the cysE gene of S. typhimurium in vitro and, more importantly, in vivo, at levels known to be able to support the synthesis of a minimum of 1.5 g of cysteine per day, indicates that the gene is well suited to use in the current sheep transgenesis programme. The use of this gene in conjunction with the cysM gene from S. typhimurium will be discussed in the following chapter.

Chapter 5

Coexpression of Cysteine Biosynthesis Genes.

Chapter 5 : Coexpression of Cysteine Biosynthesis Genes.

5.1 Introduction.

A feature central to the transgenesis programme described in this thesis is that the two genes required for cysteine biosynthesis must be coordinately expressed and inherited as a single, bifunctional unit of the sheep genome. The instances of dual-gene trangenesis are rare. To achieve this aim the desired genes must be linked in one DNA construct, such that each is independent of the other with respect to its promoter and other processing sequences. Also, to achieve coordinate expression, each gene should be subject to equivalent constraints following integration into the genome. If expression of the genes in a pair is reasonably well matched to begin with, integration into the same genomic location should afford a degree of "coordinate" expression (though strictly, this might only be referred to as coexpression). Whilst this seems the most logical approach to the transgenesis being attempted here, alternative modes of presentation of transgenes to the genome are possible and will be discussed in Chapter 6. Nonetheless, there is an obvious advantage in using genes which are linked in a non-segregating pair in our attempt to establish a permanent heritable trait.

In the preceeding chapters, a description of analyses undertaken to assess the suitability of the microbial cysteine biosynthesis genes for use in the sheep transgenesis programme was given. Although the results of these experiments have ruled out the use of the yeast gene, MET25, they have found the cysE and cysM genes from S. typhimurium to be well suited to use in this programme. The cysE and cysM genes have been successfully expressed in vitro and shown to yield enzymically active protein products which are able to function efficiently as independent species. In addition, when linked to the RSVLTR promoter, the cysE gene was expressed in vivo in a tissue-specific manner in transgenic mice and the protein product was active in many tissues. As expected, high levels of SAT activity were measured in the protein extracts prepared from tissues of mesodermal origin, such as cardiac and skeletal muscle, and various connective tissues. Expression of the RSVLTR-promoted cysM gene in these same tissues of

transgenic mice yielded low, but detectable levels of OAS sulphydrylase-B activity (A.V. Sivaprasad, unpublished data). It is known that no physical association between the cysE and cysM gene products (i.e the SAT and OAS sulphydrylase-B enzymes) occurs or is otherwise required for enzyme activity (Kredich, 1987). Hence it is reasonable to expect that expression of the genes together in a single cell should provide that cell with the capacity to synthesize cysteine from serine.

The remainder of the present chapter describes experiments designed to achieve this in vitro (permanent cell lines) and preliminary experiments aimed at coexpression to produce the complete cysteine biosynthesis pathway in vivo (transgenic mice and sheep).

Results.

5.2 Detection of expression of the cysteine biosynthesis genes.

In order to assess the success of expression of the gene combination, it would ultimately be desirable to demonstrate cysteine synthesis in intact cells. However, a limitation to the in situ assessment of enzyme production during in vitro and in vivo experiments is caused by lack of availability of the substrate sulphide, which is not found in mammalian cells. For this reason, the level of gene expression in these circumstances was initially assessed by assay of enzyme activity in protein extracts made from appropriate cell cultures and animal tissues. The cysteine synthesis assay employed was a combination of the SAT and OAS sulphydrylase assays previously used and is described in Chapter 2. Protein extracts from yeast (strain GP001), Salmonella (strain LT2) and combinations of the extracts E.1 and M.1, from the cysE and cysM permanent CHO cell lines, respectively, were tested for their capacity to synthesize cysteine. The outcome of one test assay, a 40 minute time course experiment in which the E.1 and M.1 CHO cell extracts were combined to provide the two required enzymes, is shown in figure 5.1. The TLC fractionation performed following this cysteine synthesis reaction demonstrates the expected shift in radioactivity from reactant to product

Figure 5.1 : TLC fractionation of cysteine synthesis time course and SAT assay samples.

The products of a cysteine synthesis time course reaction using a combination of 10 μg of protein from extracts made from the cysE permanent CHO cell line pLT-E.1 (E.1) and cysM permanent cell line pLT-M.1 (M.1) and 30 μM ^{14}C-serine (in 1.0 mM non-radioactive serine) as the substrate were fractionated by ascending TLC in the Checchi solvent. Time samples were taken at 0.5, 1.5, 2.5, 5, 10, 20 and 40 minutes (see the right side of the figure). The first two lanes on the far left of the autoradiogram contain the reference compounds serine, OAS and NAS, which were co-fractionated with the test samples. The third lane contains a control cysteine synthesis reaction in which the E.1 and M.1 extracts were combined and assayed for 60 minutes. In the remaining lanes, the effect of pH on the cysteine synthesis reaction and the SAT reaction (marked □) is shown. The distance of migration of the reaction products serine, OAS, NAS and cysteine (4-pyridyl-ethyl-cysteine; labelled cys-4VP) is shown The chromatogram was autoradiographed for 3 days at room temperature.
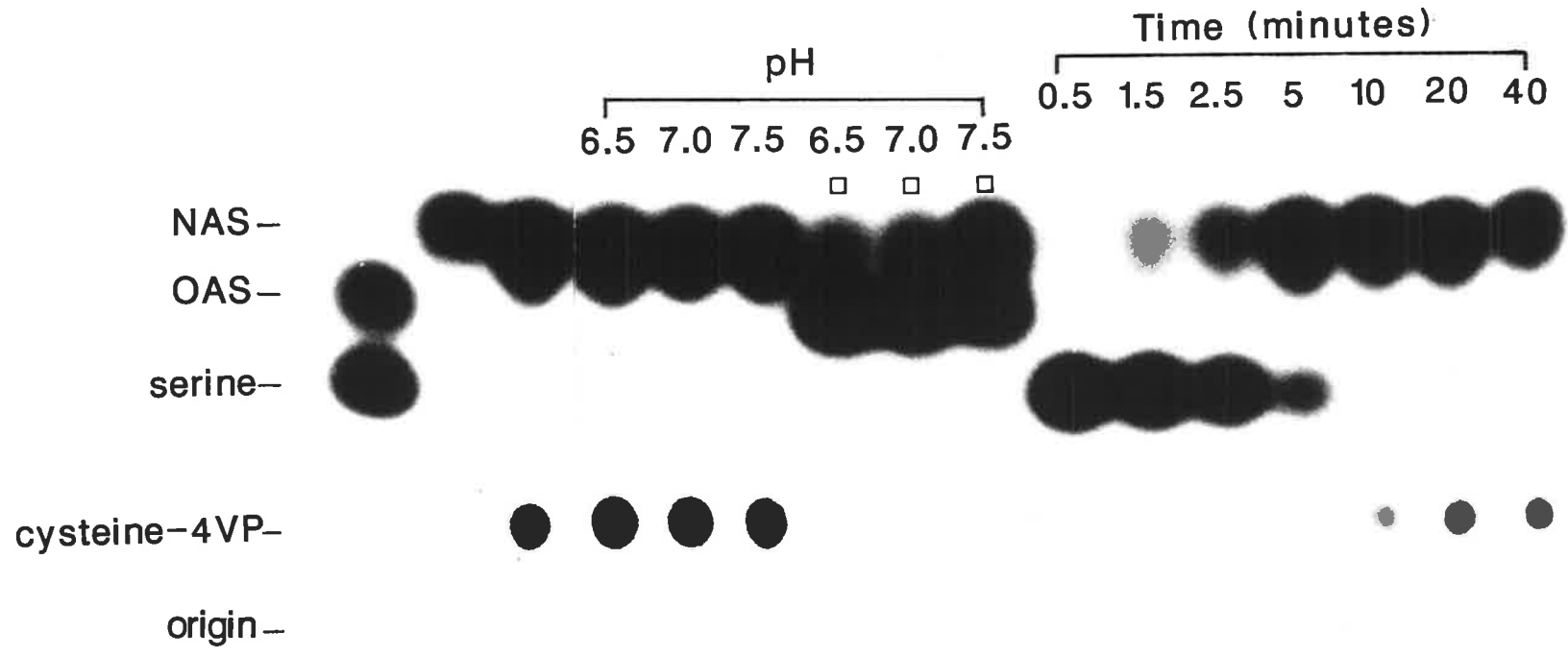
species. First serine (Rf = 0.22) is converted to OAS (Rf = 0.30) then this is converted to cysteine, which is alkylated by 4-vinylpyridine to make cysteine-4VP (Rf = 0.1).

One problem of the cysteine biosynthesis assay is the propensity of the O-acetylserine (OAS) formed in the first step of the reaction to be non-enzymically converted to N-acetylserine (NAS; Rf = 0.36) by a shift of the acetyl group which occurs under conditions where the pH is about 6.5 or greater. In some SAT assays carried out at 37 ^{0}C for 60 minutes and at pH 7.5, all of the OAS produced has been observed to be converted to NAS (data not shown). This is expected from the report that in a solution at pH 7.6 and room temperature, this acyl migration can proceed at a rate as high as 1 % per minute, increasing with alkalinity (Flavin and Slaughter, 1965). To avoid this and allow the second enzyme access to more OAS, the cysteine biosynthesis assays described in the current work were carried out at pH 6.5. A comparison of the effect of pH on enzyme function and on conversion of OAS to NAS is shown in figure 5.1 where it is clear that over the range of pH considered (from 6.5 to 7.5) there is no significant difference in enzyme activity in either the SAT or cysteine synthesis reactions, but that at the lower pH of 6.5, only a small proportion of the OAS synthesized from serine is lost to NAS by acyl migration. This is more obvious in the SAT reactions, where less compression of the reactants and products in the TLC fractionation occurs. It should be noted that such compression is a normal feature of the TLC fractionations when high salt concentrations are present, and will be apparent in most of the fractionations of cysteine synthesis reactions. Consequently, the OAS and NAS spots in fractionated cysteine synthesis assay samples will be less discrete.

Measurement of the level of gene expression by assay of enzyme activity was complemented by quantitative analysis of transcript levels in total RNA prepared from the same cell culture and sheep tissue samples (see sections 5.4.2 and 5.5.2.2). Using suitable complementary RNA probes, the RNA protection method described in Chapter 2 (Melton et al., 1984; Krieg and Melton, 1987) was used to compare the levels of cysE and cysM transcripts present in RNA samples, with a view to determining whether transcripts derived from each gene in the linked-gene construct (described below) were equally represented in the mRNA.

5.3 Construction of a linked-gene eukaryotic expression plasmid.

The requirement for inheritance of the cysteine biosynthesis genes as a non-segregating pair in transgenic animals dictated the construction of a eukaryotic expression plasmid in which the cysE and cysM genes were physically linked but functionally independent. The ability to isolate a convenient fragment for transgenesis also had to be considered during construction of this plasmid. In order to retain the level of expression achieved in the previous work, the structure of the hybrid cysE and cysM genes was preserved (i.e. transcription of each of these genes was governed by the RSVLTR promoter and post-transcriptional processing by the hGH 3' end sequences; these sequence elements are as previously described in the plasmid constructs pLT-MET, pLT-E and pCS942; see figures 3.9 b, 4.2 b and 5.2). Another major consideration in the linkage of the two genes was whether to arrange them such that transcription from their respective promoters was divergent, convergent or in tandem, and in any case so as to minimise transcriptional interference which could arise by the simultaneous operation of promoters situated close to (or even at some distance from) each other (Proudfoot, 1986). Considering that each of these types of gene arrangement has been observed in the gene clusters found in nature (e.g. in the histone and keratin multi-gene families and the globin gene locus, and in relatively small genomes such as those of the SV40 virus and bacteriophages Lamda and 186) and that a high level of transgene expression is possible following the integration of fragments in a head to tail arrangement in multicopy loci, the last of these three options, with the genes cloned in tandem, was chosen as a model for the construction of a linked-gene construct.

Construction of a linked-gene plasmid involved cloning of the entire cysM hybrid gene as a 3.3 kb fragment (isolated from pCS942) into the ScaI site of the construct pLT-E. The details of construction of this new plasmid pLT-EM are given in the legend to figure 5.2. When the cysM gene was inserted in tandem with the cysE gene, the ScaI site at the 3' end of the fragment was preserved. This allowed excision of a 6.4 kb SalI / ScaI fragment for transgenesis, containing both of the genes and all of their processing sequences. In this tandem, head to tail arrangement, the cap site of the RSVLTR promoter belonging to the

Figure 5.2 : Construction of linked-gene expression vector pLT-EM

A 3.3 kb SalI / ScaI fragment was isolated from vector pCS942 (A.V. Sivaprasad, unpublished; see (i)) , treated with Klenow enzyme and cloned by blunt-end ligation into ScaI-digested / CIP-treated pLT-E (ii) to create pLT-EM (iii). During the cloning, the ScaI site was preserved. and this enabled the excision of a 6.4 kb SalI / ScaI linked-gene fragment which could be used for transgenesis in mice and sheep. A linear map of the 6.4 kb fragment in (iv) shows the 3.1 kb and 3.3 kb unit lengths of the respective cysE and cysM gene segments (each including its own RSVLTR promoter and hGH 3'-end processing sequences) and the distance between the EcoRI sites is 3.1 kb. If the 6.4 kb linked-gene fragment is integrated into a host genome in a head to tail array, succesive 3.1 kb and 3.3 kb fragments wiil be excised by EcoRI, along with two fragments of unknown size containing some plasmid DNA and genomic DNA flanking the transgene locus.

(i) **pCS942** 7.81 Kb — SalI, EcoRI, BamHI, RSVLTR, SV40, cysM, hGH, neo, amp, ScaI, pBR ori

(ii) **pLT-E** 7.49 Kb — SalI, EcoRI, ATG, RSVLTR, SV40, cysE, TGA, hGH, neo, amp, ScaI, pBR ori

(iii) **pLT-EM** 10.85 Kb — SalI, EcoRI, RSVLTR, cysE, SV40, hGH, neo, SalI/ScaI, RSVLTR, EcoRI, cysM, hGH, pBR ori, amp, ScaI

(iv) SalI, EcoRI, RSVLTR, cysE, hGH, SalI/ScaI, EcoRI, RSVLTR, cysM, hGH, ScaI, 3.1 kb, 3.3 kb
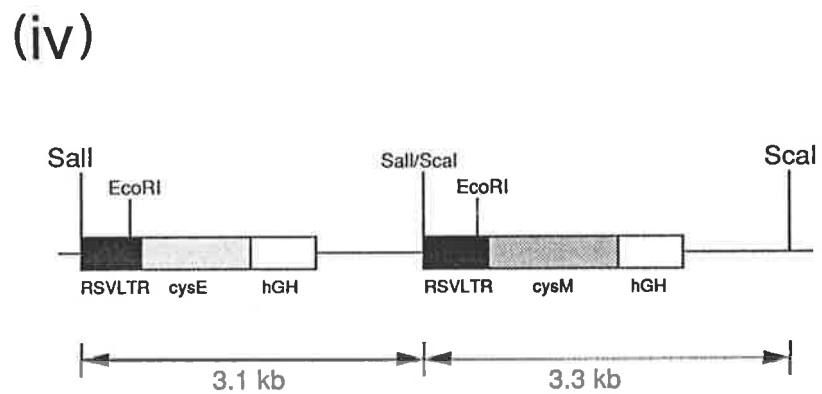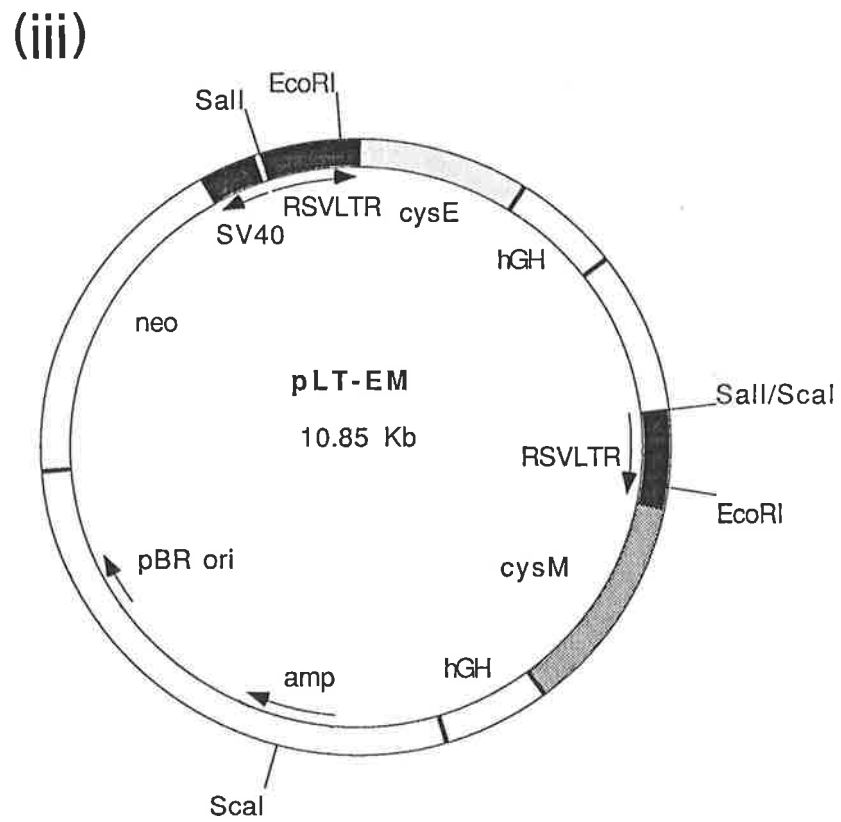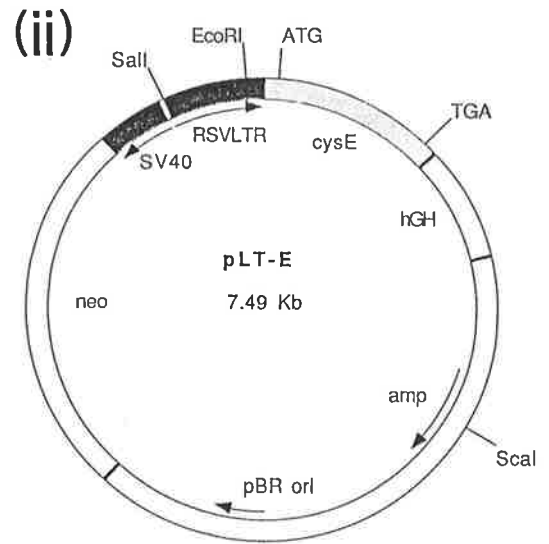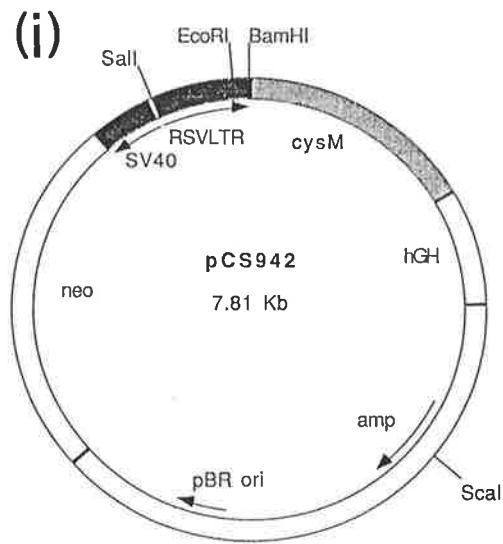
cysM gene is separated from the polyadenylation signal (from hGH) of the cysE
gene by approximately 1.8 kb and the promoters of the two genes by 3.1 kb. The
proximity and directionality given to the genes in this case is not unlike that seen
in some of the gene clusters described previously. Retention of the neo^r gene
marker was essential in use of the plasmid to produce permanent CHO cell lines.

5.4 Coexpression of the cysE and cysM genes in vitro.

5.4.1 Coexpression in transient cell cultures.

A study of the coexpression of the cysE and cysM genes in the linked-
gene construct was begun by transient transfection of CHO and COS-1 cells with
the plasmid pLT-EM. Electroporation was employed to introduce supercoiled
plasmid DNA into these cells. When a portion of the protein extracts made from
these transfected cells was assayed in a cysteine synthesis assay, conversion of
the ^{14}C-serine to OAS was observed in each case (figure 5.3). Unlike the CHO
cell extract, where even the conversion of serine to OAS was incomplete
(approximately 50 %; figure 5.3 a) the COS-1 protein extract converted all of the
input serine to OAS, and in addition, some cysteine was produced (figure 5.3 b).
This difference between cell lines could be expected, considering the presumed ability of
plasmid pLT-EM (containing an SV40 ori sequence) to replicate in the COS-1
cells with the aid of the large T-antigen produced there. The protein extracts made
from the cells of negative control CHO and COS-1 cell transfections (no DNA
was included in the electroporations) had no such enzyme activities.

5.4.2 Coexpression in permanent cell cultures.

The transient expression experiment immediately demonstrated that each
gene in the construct was being accurately expressed. However, the level of
activity of each enzyme was comparatively low, especially in the case of the CHO
cell protein extract. In addition, this experiment could not distinguish whether
each of the genes of a particular linked-gene plasmid molecule were being
expressed as opposed to the cysE gene of one plasmid molecule and the cysM
gene of another, in separate cells. A solution to both of these problems was
sought in the creation of permanent CHO cell lines by transfection with DNA of

Figure 5.3 : Cysteine synthesis assay analysis of protein extracts from mammalian cells transfected with construct pLT-EM.

Following transfection of CHO (a) and COS-1 (b) cells with construct pLT-EM by electroporation (a negative control included no DNA) cells were maintained in culture for 48 hours, then harvested for the preparation of protein extracts. Cysteine synthesis assays were performed using approximately 10 µg of protein and 30 µM ^{14}C-serine in the assay, the reaction products fractionated by ascending TLC in the Checchi solvent and the chromatograms autoradiographed for 3 days at room temperature. In (a) the positive controls are GP001 yeast extract (+) and a mixture of the E.1 and M.1 extracts (E.1 + M.1) and in (b) only the latter was used. The reference compounds serine (S) OAS (O) and NAS (N) were co-fractionated with test samples in each case and the position of expected reaction products is indicated on the left side of the figure.

a.

NAS-
OAS-
serine-

cys-4VP-
origin-

S O N + E.1+M.1 No DNA pLT-EM

CHO

b.

NAS-
OAS-
serine-

cys-4VP-
origin-

S O N E.1+M.1 No DNA pLT-EM

COS-1
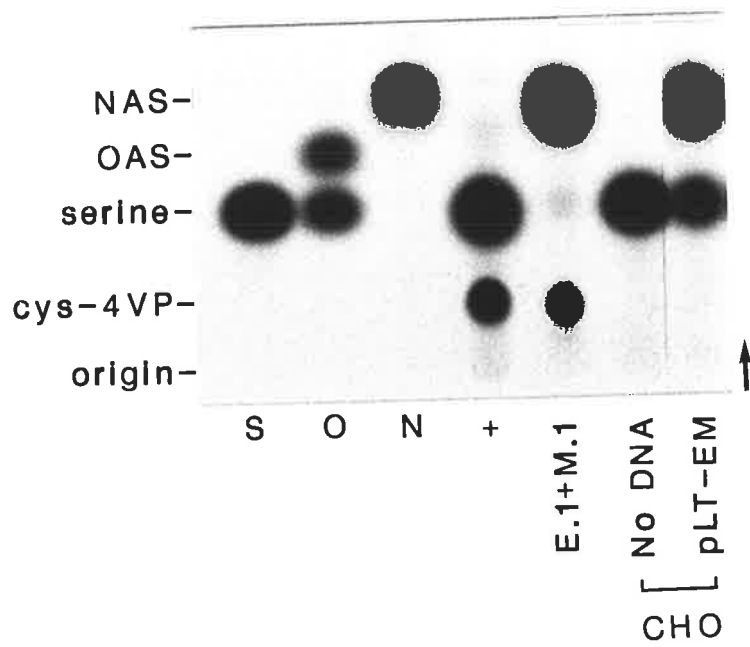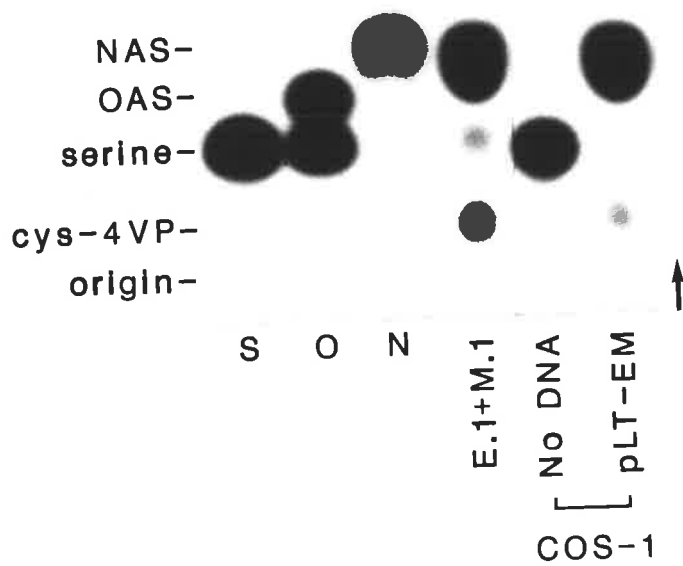
the plasmid pLT-EM linearised at the SalI site. Ideally, expression of each gene in such a cloned cell line containing only a single copy of the linked-gene construct would indicate successful coexpression of the genes when in this arrangement and would more accurately predict the level of coexpression achievable in vivo. Resistance to 350 µg/ml of the neomycin analogue, G418, was used to select cells expressing high levels of the aminoglycoside phosphotransferase enzyme from stably integrated copies of the plasmid constructs. Using this technique, six G418-resistant cell lines (pLT-EM.1 - .6) were selected and maintained in medium containing G418 for many months to firmly establish their character.

5.4.2.1   Initial characterisation of permanent cell lines.

When 20 µg of protein in extracts made from these cell lines was assayed for cysteine synthesis, all but one, pLT-EM.1, displayed both of the desired enzyme activities (figure 5.4 a) albeit at markedly varying levels. When quantitated, the ability to convert serine to cysteine (via OAS) varied from 0.1 % conversion in line pLT-EM.3 to 27.3 % conversion in line pLT-EM.4. As in the case of a negative control cell line, pRSV.2 (see chapter 3) the line pLT-EM.1 was found to possess neither of these activities.

When a dot-blot hybridization analysis of genomic DNA prepared from these cell lines was performed using the 6.4 kb SalI / ScaI fragment from plasmid pLT-EM as a probe (figure 5.4 b), pLT-EM.1 was seen to contain only one copy of the 11 kilobase plasmid. The genomes of the other cloned lines contained 2 - 5 copies of the plasmid DNA. This was confirmed by Southern transfer analysis of EcoRI-digested genomic DNA prepared from each of the cell lines pLT-EM.1 - pLT-EM.6 (referred to as lines EM.1 - EM.6) and the negative control pRSV.2 (figure 5.4 c). In the case of line EM.1, the Southern pattern is unclear, and may indicate integration of a sub-fragment of the plasmid. Only a single high molecular weight band greater than 7.8 kb in size is evident in the EcoRI-digested DNA and another even higher molecular weight species seems to be present. If this is the case, then only one of the EcoRI sites of the integrated plasmid molecule has been cleaved (i.e the probe did not detect the 3.1 kb EcoRI fragment which is expected in a digest of the linked gene construct pLT-EM; see figure 5.2). This result is similar to that seen in one of the integrated transgene copies of the cysE transgenic E-16 mouse

Figure 5.4 : Characterisation of the G418-resistant pLT-EM CHO cell lines.

(a) : Cysteine synthesis assay of protein extracts from the pLT-EM cell lines.

20 µg of protein from each of the cell lines pLT-EM 1 - pLT-EM.6, and negative control cell line pRSV.2 (see section 3.4.2) was assayed to detect activity of the cysE and cysM gene products in a cysteine synthesis assay. The reaction products were fractionated by ascending TLC along with appropriate reference compounds and the chromatogram autoradiographed for 3 days at room temperature. The migration distance of expected reaction products is indicated on the right side of the figure.

(b) : Dot-blot hybridisation analysis of genomic DNA from the pLT-EM cell lines.

5 µg of genomic DNA from each of the pLT-EM cell lines (.1 - .6) was transferred to Genescreen membrane and hybridised with the 6.4 kb SalI / ScaI linked-gene fragment which was ^{32}P-radiolabelled by Nick-translation. A copy control series made using appropriate amounts of the plasmid pLT-EM (diluted into 5 µg of genomic DNA from a non-transgenic mouse) was also transferred to the membrane. After hybridisation, the filter was washed at 0.2 x SSC / 65 ^{0}C and autoradiography carried out at -80 ^{0}C for 16 hours.

(c) : Southern transfer hybridisation analysis of EcoRI-digested genomic DNA from the pLT-EM cell lines.

10 µg of EcoRI-digested genomic DNA from the pLT-EM cell lines (EM.1 - EM.6) and line pRSV.2 was analysed by Southern transfer hybridisation using Zetaprobe membrane. The 6.4 kb SalI / ScaI fragment was radiolabelled as described in (b) and hybridised to the membrane bound DNA, then the filter washed and autoradiographed as in (b). The size (kb) and position of marker DNA fragments (M; HindIII-digested Lambda DNA) is indicated on the left side of the figure.

a.

- NAS
- OAS
- serine
- cys-4VP
- origin

^{14}C-SER ^{14}C-OAS ^{14}C-NAS pRSV.2 EM.1 EM.2 EM.3 EM.4 EM.5 EM.6

b.

.1 .2 .3 .4 .5 .6

- EM cell lines
- Copy number

0 0.5 1 2 5 10 25

c.

M pRSV.2 EM.1 EM.2 EM.3 EM.4 EM.5 EM.6

k.b.

23.1 -
9.4 -
6.6 -
4.4 -

2.3 -
2.0 -

0.56 -

◄ 7.8 kb

◄ 3.1 kb
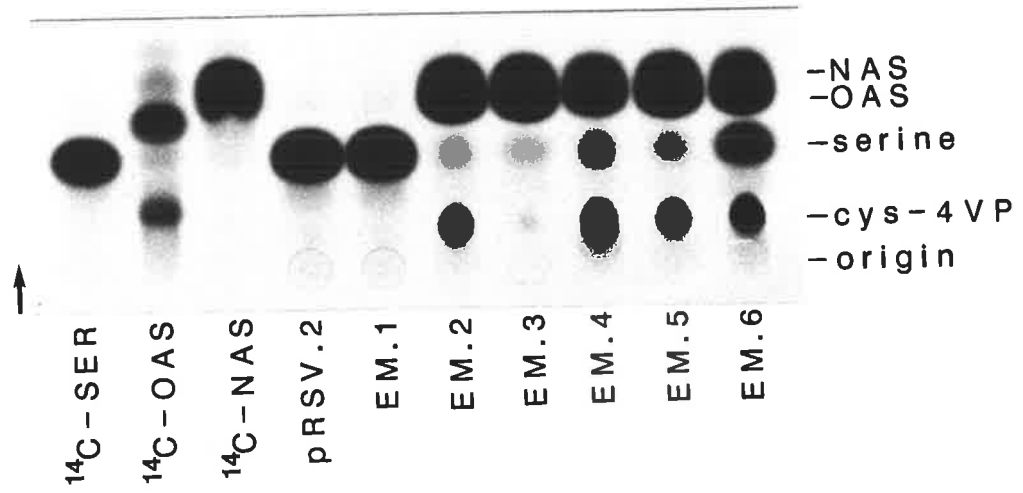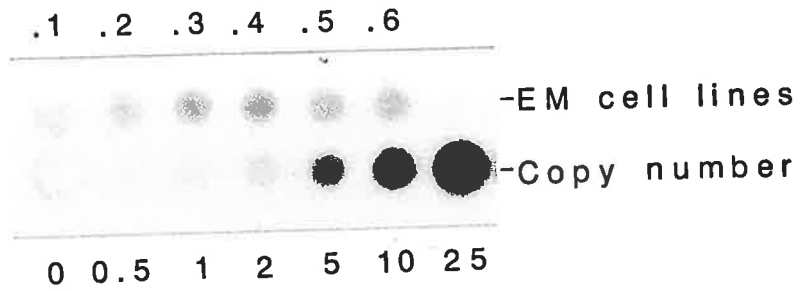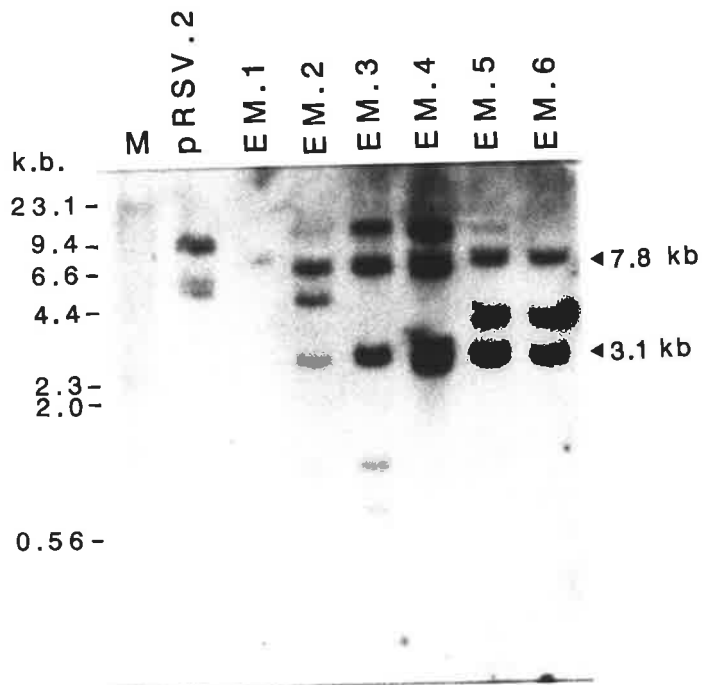
line (see section 4.3 c) and proposed for other F_0 cysE transgenic mice where lack of expression of the integrated transgenes was accompanied by lack of cleavage at the expected EcoRI sites.

In DNA of cell lines EM.2 - EM.6, the band patterns were indicative of insertion of multiple copies of the plasmid. The appearence of bands at 3.1 kb and 7.8 kb indicated the presence of at least two intact copies of the plasmid in a head to tail arrangement in the genome of each of these cell lines. In addition, in the DNA of cell line EM.3, a band detected by the probe at about 900 bp indicated a head to head arrangement of at least two of the plasmids in the genome of this line with the RSVLTR-promoted cysE genes of adjacent integrated plasmids oriented for divergent transcription. This band was also visible in the DNA of cell line EM.2 in a longer autoradiographic exposure of the Southern filter (data not shown). Other EcoRI fragments detected in this Southern analysis are consistent with expected end-fragments containing either the 5' or 3' termini of the integrated plasmids connected to genomic flanking sequences. Assuming this, the similarity of the sizes of these additional fragments in the genomes of cell lines EM.5 and EM.6 suggested that they were possibly separate isolates of the same clone or that the integration had occurred at the same genomic site in the case of each clone. (As stated previously in section 4.4, a degree of site specific integration has been attributed to the RSVLTR; Shih et al., 1988.) Finally, the band pattern of the negative control cell line pRSV.2 is indicative of tandem genomic integration of two copies of the base vector pRSVΔH. Digestion of the DNA of this cell line with EcoRI has produced a unit length 6.5 kb fragment and two end fragments containing plasmid sequences and genomic flanking DNA.

5.4.2.2   Analysis of coexpression in the permanent cell lines.

To begin analysis of the coexpression of the cysE and cysM genes in the linked-gene permanent CHO cell lines, the levels of the respective transcripts in total cytoplasmic RNA made from each cell line were compared using RNA protection assays. The 331 base ^{32}P-labelled antisense probe used initially was derived from the plasmid clone pLTRE-B (figure 5.5 a) and prepared by SP6 transcription and PAGE purification as described in Chapter 2. This probe is able to hybridize and protect approximately 285 bases of cysE mRNA, including 40

Figure 5.5 : Construction of RNA protection probe transcription vectors pLTRE-B and pLTRM-B.

(a) : A 336 bp EcoRI / AluI fragment was isolated from pLT-E and cloned directionally into HincII / EcoRI-digested, CIP treated pGEM-1 to create pLTRE-B. After linearisation with EcoRI, synthesis of SP6 transcripts in the presence of ^{32}P-rUTP was performed to generate a complementary antisense probe specific for cysE transcripts arising from the RSVLTR promoter in pLT-EM, and which protects approximately 285 bases of cysE RNA. This probe also shares 40 bp of sequence complementary to the cysM gene transcripts produced by pLT-EM.

(b) : A 301 bp EcoRI / RsaI fragment was isolated from pLT-EM and cloned directionally into HincII / EcoRI-digested, CIP-treated pGEM-1 to create pLTRM-B. Linearisation with EcoRI and synthesis of SP6 transcripts in the presence of ^{32}P-rUTP generated a complementary antisense probe specific for cysM transcripts arising from the RSVLTR promoter in pLT-EM, and which protects approximately 250 bases of cysM RNA. This probe also shares 40 bp of sequence complementary to the cysE gene transcripts produced by pLT-EM.

* The RSVLTR transcription initiation site (51 bp from the EcoRI site) is marked (⤙).

a.

AluI — HincII

245 b
cysE

40 b   51 b
RSVLTR

EcoRI — EcoRI

SP6   T7

pLTRE-B
3.2 Kb

pBR ori

amp

ScaI

b.

RsaI — HincII

210 b
cysM

X. B-g

40 b   51 b
RSVLTR

EcoRI — EcoRI

SP6   T7

pLTRM-B
3.14 Kb

pBR ori

amp

ScaI
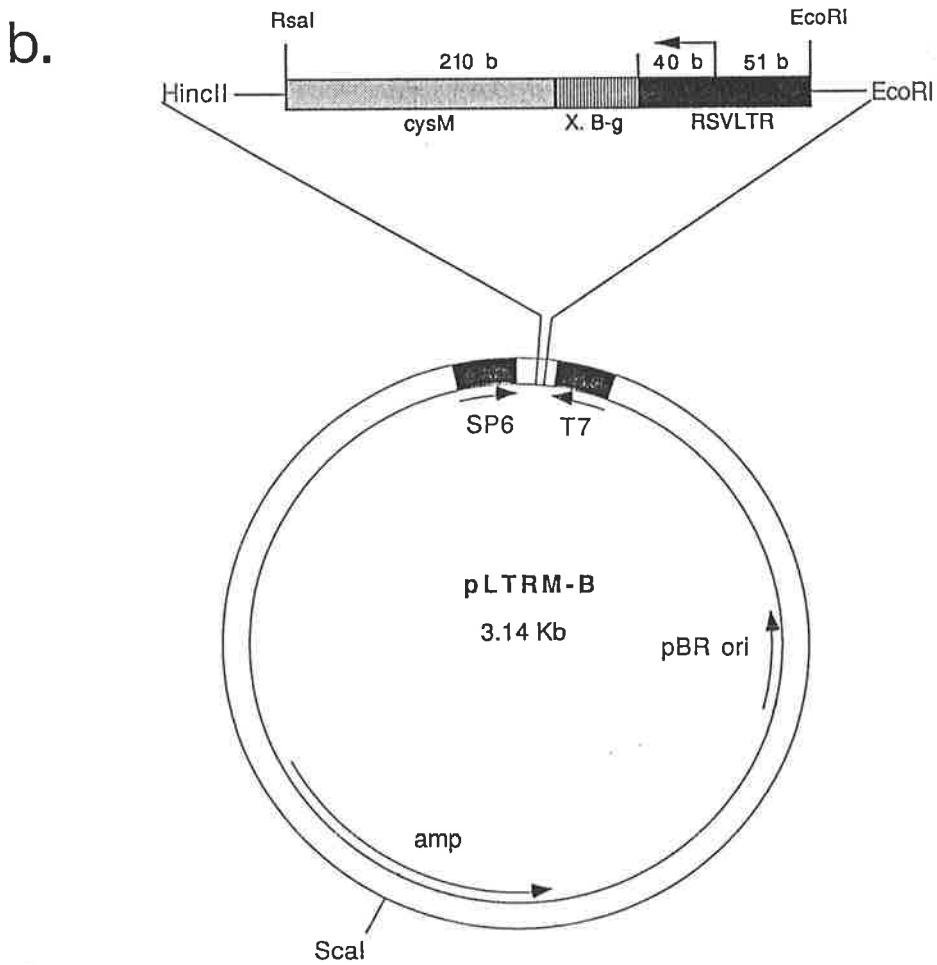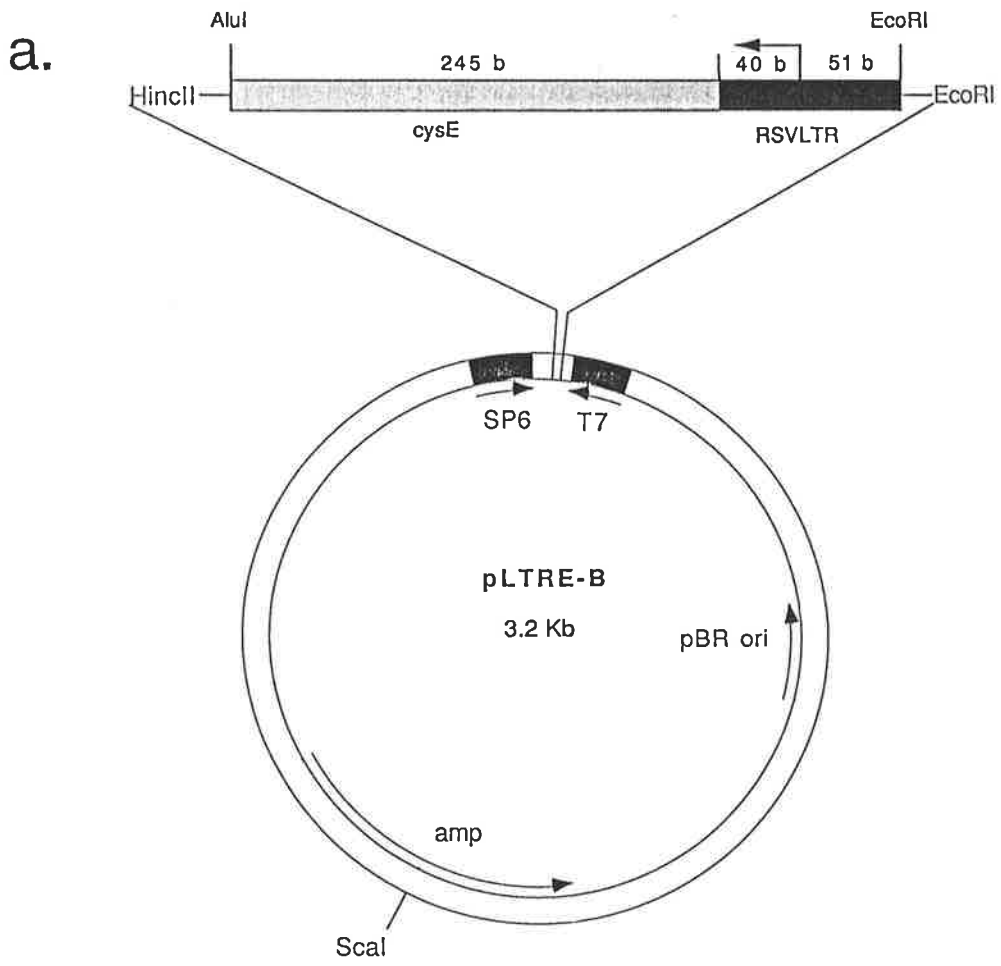
bases of RSVLTR sequence 3' to the cap site and 245 bases from the 5'-untranslated and coding regions of the cysE gene. The 40 bases of RSVLTR sequence is common to both the cysE and cysM transcripts, hence the cRNA probe is able to detect both mRNA species in the same sample. With this in mind, RNA protection analyses of 2.5 μg of cytoplasmic RNA prepared from cell lines EM.2 - EM.5 and the negative control cell line pRSV.2 were performed (figure 5.5 c). Two main features immediately obvious from these analyses are the difference in the general level of expression of the individual genes between cell lines and the difference between the ratio of expression of cysE : cysM within a particular cell line. For example, the level of cysE gene expression is high in EM.4 but extremely low in EM.6 and only moderate in the other lines. Secondly, the level of cysE mRNA transcripts in line EM.6 is very low compared to that of the cysM gene transcripts in this same cell line, whereas in line EM.3 the converse is true with cysE being moderately expressed and cysM being very poorly expressed.

To illustrate the different levels of cysE and cysM transcripts in individual cell lines more clearly, this experiment was repeated twice more, once using the antisense probe derived from pLTRE-B as before, and in the second instance using a probe derived from clone pLTRM-B (figure 5.5 b). Converse to the pLTRE-B probe, this new probe could protect approximately 250 bases of the cysM transcript and otherwise shared 40 bases of the RSVLTR sequence in common with the cysE transcripts.(The structures of the RNA protection clones pLTRE-B and pLTRM-B are compared in figure 5.5 a and b.) The results of these analyses are shown in figure 5.5 d (the cysM protection product being 247 bases). Where possible, the variability in the level of expression of the cysE and cysM genes in the linked-gene construct in these indiviual cell lines was confirmed quantitatively by scintillation counting of ^{32}P-labelled protected bands (from the 500 pg positive control [see E in figure 5.5 c] and test samples) excised from the gels following autoradiography, and by densitometry of the autoradiograms produced. In the calculation of cysE and cysM transcript levels and ratios, the size difference between cysE and cysM protected fragments, and the difference betweeen the sizes of these and their respective positive control protected transcripts were taken into account. A ratio of the levels of cysE

Figure 5.5 (c) : RNA protection analysis of cysE and cysM gene transcription in cell lines pLT-EM.1 - pLT-EM.6.

A 2.5 μg sample of cytoplasmic RNA isolated from cell lines EM.1 - EM.6 and from cell line pRSV.2 was analysed by RNA protection analysis using the pLTRE-B antisense probe (radiolabelled as described in figure 5.5 a). Half of each sample was fractionated in a denaturing 6 % (19:1) Urea-PAGE system alongside probe control samples (Y+ and Y-; see Methods section 2.2.2) and a cysE positive control (E; 500 pg of cysE transcript from pGEM-cysE in 2.5 μg of yeast tRNA) which gives a protected fragment of approximately 220 bases in size. To allow sizing of the protected fragments, ^{32}P-end-labelled HpaII-digested pUC19 DNA markers were co-fractionated with the RNA protection samples. The two marker and control tracks were autoradiographed for 6 hours at room temperature and the experimental tracks for 48 hours at -80 ^{0}C with intensification. The position of the 285 b cysE and 40 b cysM protected fragments is indicated in the figure.

C.
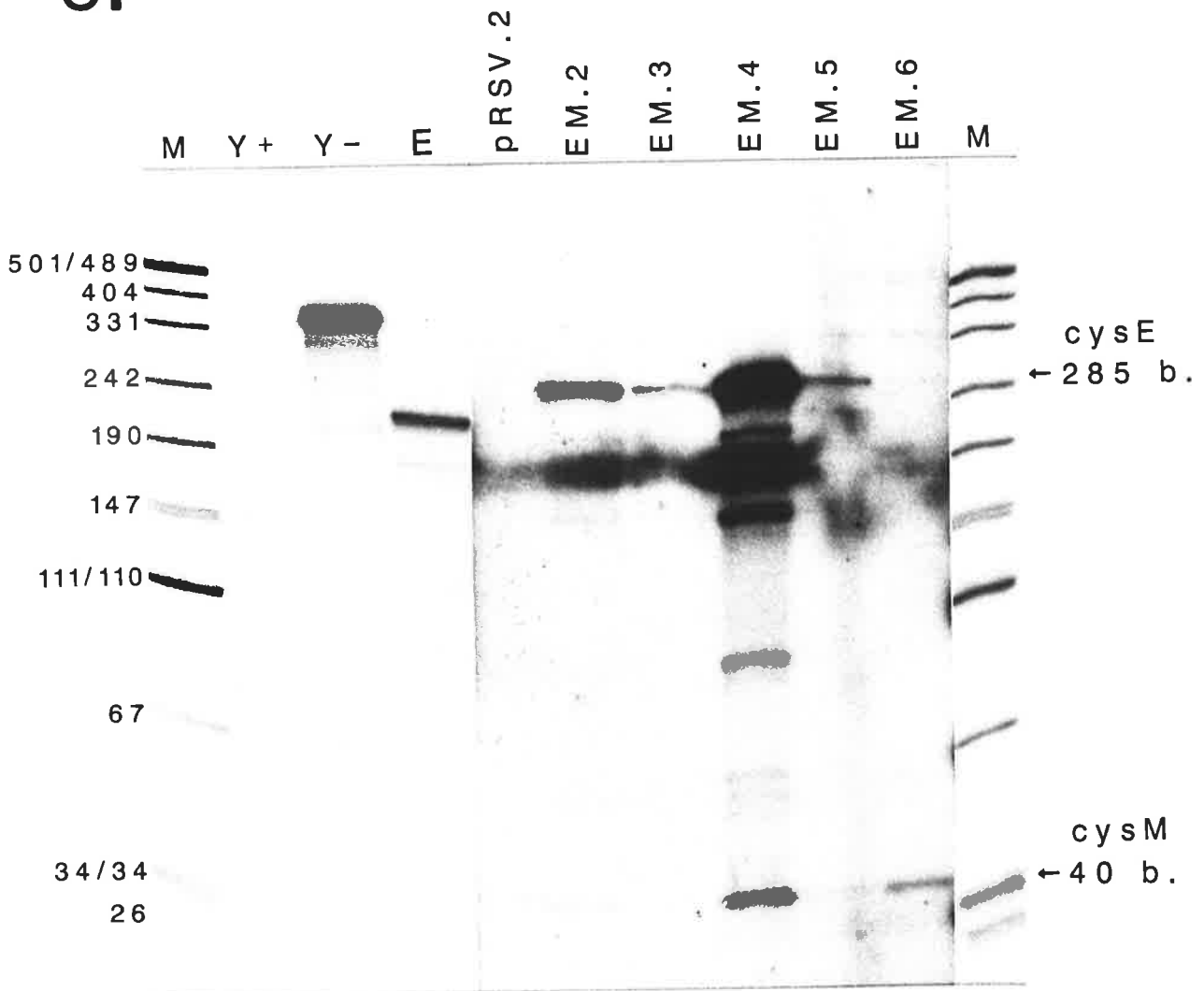
Figure 5.5 (d) : RNA protection analysis of cysE and cysM gene transcription in cell lines pLT-EM.1 - pLT-EM.6.

The RNA protection analysis described in figure 5.5 (c) was repeated using radiolabelled antisense probes from both the pLTRE-B (cysE-specific probe) and pLTRM-B (cysM-specific probe) clones, and half of each of the samples (only one-tenth of the probe control reactions Y+ and Y- was loaded) processed as previously described. In the top panel, the result of protection using the cysE-specific probe is shown and the position of the 285 b protected fragments indicated The positive controls are 50 and 500 pg of cysE sense transcript derived from pGEM-cysE. In the lower panel, RNA protection using the cysM-specific probe is shown and the position of the 247 b protected fragments indicated. In this case, the positive controls are 50 and 500 pg of cysM sense transcript derived from pCS920 (A.Sivaprasad; unpublished).The protected fragment in this case is approximately 210 bases in size. Autoradiography of each of the gels was carried out for 28 hours at -80 ^{0}C with intensification.

d.

M	Y+	Y−	50pg	500pg	pRSV.2	EM.2	EM.3	EM.4	EM.5	EM.6	M
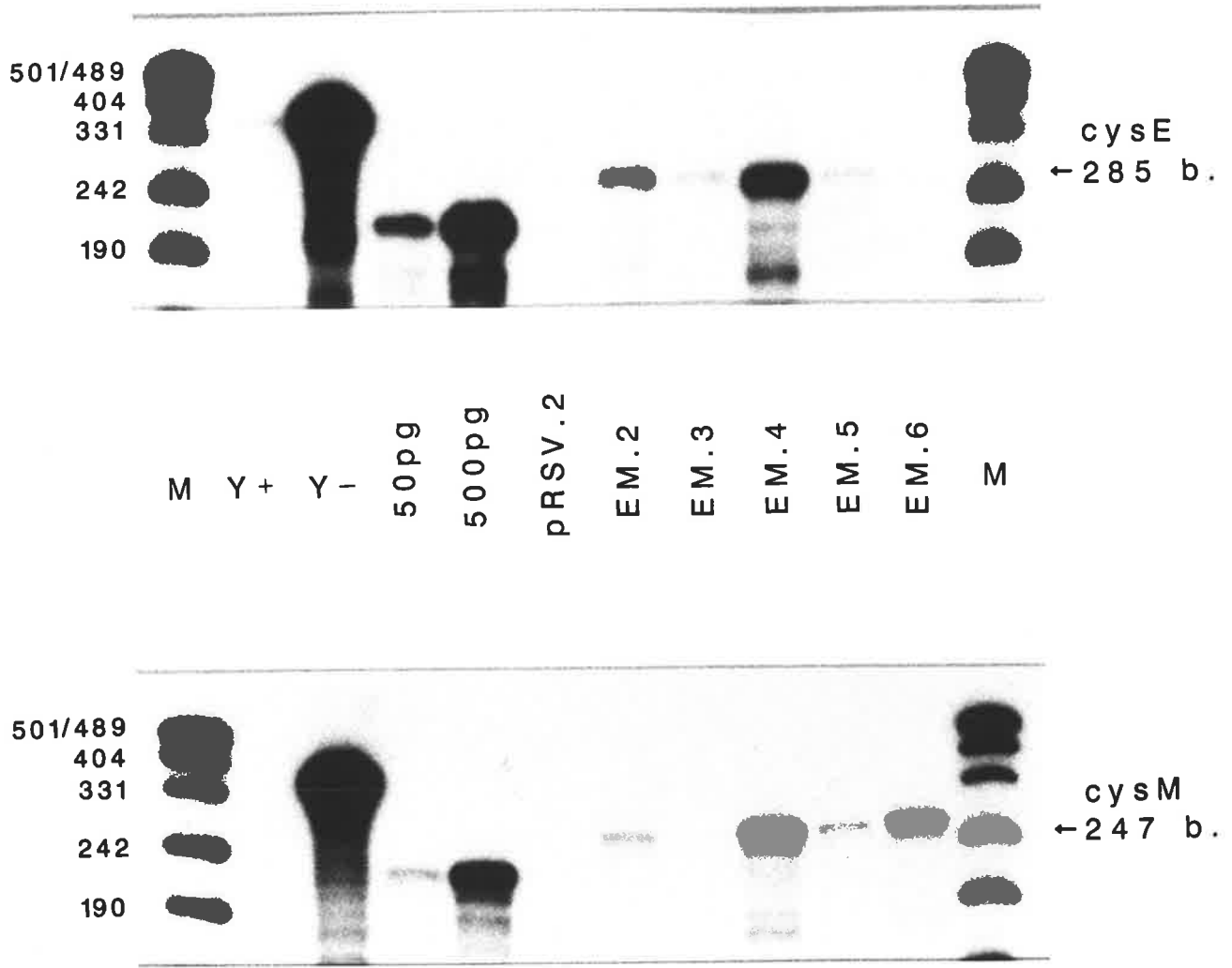
cysE
← 285 b.

cysM
← 247 b.

transcripts present in the RNA of cell lines pLT-EM.2 - pLT-EM.6 are given in figure 5.6 b (table 2).

Considering only the example of cell line EM.4, where the ratio of the level of cysE : cysM transcript is of the order of 2.8 : 1.0, there are hundreds of picograms of each of these transcripts per 2.5 µg of RNA analysed (in fact about 700 pg and 250 pg respectively). Assuming that 10 µg of total cellular RNA represents the RNA from approximately 1 million eukaryotic cells (Lee and Costlow, 1987) an approximation of the number of copies of the transcripts per cell can be made. In the case of cell line EM.4, assuming transcript sizes of 1.2 kb for cysE and 1.4 kb for cysM, this suggests some 4000 cysE and 1200 cysM transcripts per cell which agrees with the level of expression of these genes as determined at the protein level (see sections 5.4.2 a and 5.4.2 c). In many cases however, the levels of transcript were sufficiently low in amount so as to cause inaccuracy in their determination. One extra fact emerging from this analysis was that cell lines pLT-EM.5 and .6 are not different isolates of the same original clone, as suggested from their Southern banding patterns in figure 5.4 c, but are distinct clones on the basis of the different levels of cysE and cysM transcript that they produce.

Regarding the level of enzyme activity associated with the RNA detected in the five EM cell lines, a crude quantitation of the SAT activity in protein extracts (made from the same cell samples from which the RNAs were derived) was performed via time course assay of enzyme activity in 20 µg of total cell protein (figure 5.6 a). The approximate level of SAT enzyme present in each of the extracts assayed is expressed as milliunits per milligram of protein in table 1, figure 5.6 b. When the ratio of SAT enzyme activity between the pLT-EM cell lines was compared with the ratio of cysE transcripts present (figure 5.6 b, table 2) those for cell lines EM-2, EM-3, EM-5 and EM-6 compared favourably in magnitude. In the case of cell line EM-4, the transcript level appeared appreciably higher than expected considering the level of SAT activity determined. However, as the time course curve for this sample indicates that most of the conversion of serine to OAS occurred in the first 10 minutes (see figure 5.6 a), it is clear that the measurement of SAT activity in cell line EM-4 was limited by substrate availability. This accounts for the variation in SAT activity and transcript levels in

Figure 5.6 Quantitation of SAT activity in cell lines pLT-EM.2 - pLT-EM.6.

(a) : A SAT assay time course analysis of 20 μg of protein extracts from pLT-EM.2 - pLT-EM.6 was performed as previously described, with time samples taken every five minutes from 0 - 60 minutes. The reaction products were fractionated by ascending TLC in the Checchi solvent and the chromatogram autoradiographed for 3 days. This was followed by densitometric analysis of the autoradiogram, and the % conversion of serine to OAS is graphed here as a function of time.

(b) : Table 1.

Determination of the specific activities of the protein extracts assayed in (a) with respect to SAT activity was carried out and these are expressed in milliunits per milligram of protein for each cell line extract.

: Table 2.

A comparison of the relative specific activities of SAT enzyme measured by assay of the protein extracts and levels of cysE transcript measured by RNA protection analyses (by densitometry of autoradiograms and scintillation counting of protected fragment bands excised from gels) for each of the five EM cell lines is shown. The ratio of SAT activity between the cell lines was expressed relative to that measured in line EM.6 and the same was done with respect to the level of cysE transcripts detected in the RNA protection analyses.

a.

SAT Assay Time Course : EM cell lines
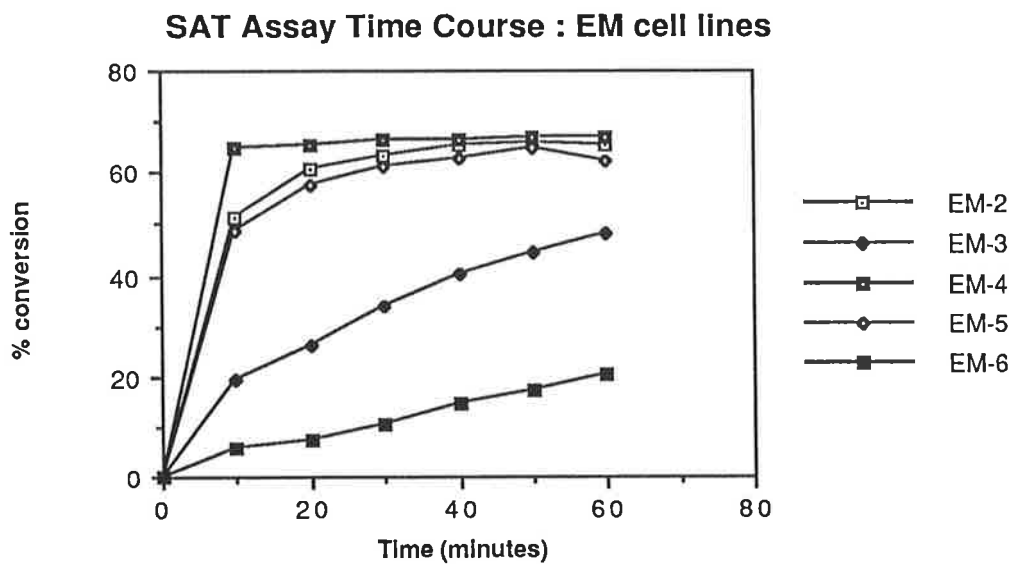
Cell line	SAT activity (milliunits/mg)
EM-2	53.5
EM-3	19.6
EM-4	66.9
EM-5	50.5
EM-6	5.9

b.

Table 1

Table 2

Cell line	Ratio of SAT activity	Ratio of cysE transcript levels
EM-2	9.0	14.8
EM-3	3.3	3.35
EM-4	11.3	104.8
EM-5	8.7	4.95
EM-6	1.0	1.0

this case and overall, the level of SAT activity measured for a cell line reflects the level of cysE transcript present. Given that the efficiency of translation of cysM transcripts derived from this construct is poorer than that of cysE (initially observed in vitro in the RRL translation system; A.V. Sivaprasad, unpublished data) it was otherwise assumed that the correlation between RNA and enzyme levels demonstrated for the cysE gene would also generally apply to the cysM gene. However, to be certain that no other factor affects production of the cysM gene product in vivo, the OAS sulphydrylase enzyme and cysM transcript levels must be separately analysed.

One facet of expression highlighted by the transcript analysis performed here was the effect of differing ratios of cysE : cysM gene products expressed in the various pLT-EM cell lines upon cysteine synthesis. The result of this is particularly evident in the initial cysteine synthesis assays (see figure 5.4 a). For example, in line pLT-EM.3 (figure 5.4 a) where the ratio of cysE : cysM expression is high, the first-step SAT reaction has proceeded to completion, but very little cysteine has been formed in the second-step sulphydrylase reaction. In contrast to this, in line pLT-EM.6 (figure 5.4 a) where the ratio of cysE : cysM expression is low, the SAT reaction has proceeded very slowly whilst the sulphydrylase reaction has produced a large quantity of the end-product cysteine from the limited OAS available. In this latter instance is a demonstration of the possible use of limitation of the level of SAT activity such that the sulphydrylase enzyme can then utilise a higher proportion of the OAS produced in the first reaction step, especially before it is lost as NAS. In addition, where the SAT enzyme is feed-back inhibited by relatively low levels (i.e. 1.1 uM) of the end-product cysteine, there may be some advantage in restricting the level of cysteine synthesis in vivo. These aspects of the cysteine synthesis reaction will be discussed further in section 5.6.

5.4.2.3 Analysis of the cysteine synthesis capacity of cell line pLT-EM.4.

The SAT and OAS sulphydrylase activities of cell line pLT-EM.4, which displays the highest level of expression of the cysE and cysM genes, were quantitated in order to measure its capacity for cysteine synthesis. A protein

extract prepared from pLT-EM.4 was tested for SAT and OAS sulphydrylase activity in a cysteine biosynthesis time course assay in which none of the substrates was limiting. For the purpose of determining potential cysteine synthesis capacity of similar cells in vivo, it was thought more relevant to examine the effect of the enzyme combination, rather than to assess enzyme activities separately. Along with the result of the TLC fractionation of the products of this time course reaction (figure 5.7 a) is a graph (figure 5.7 b) which reveals a maximum rate of cysteine synthesis of 35 nanomoles per minute per 20 µg of protein. From the autoradiogram, it is evident that the specific activity of OAS sulphydrylase-B in this extract is lower than the SAT activity. This is consistent with the finding of a lower level of cysM transcripts than cysE transcripts in this cell line, and may also reflect the lower affinity of the former for its substrates and the poorer translation efficiency of the cysM transcript (see section 5.6.1 and Chapter 6).

This result clearly shows that coexpression of the cysteine biosynthesis genes of S. typhimurium in vitro in mammalian cells has conferred the potential for cysteine synthesis on these cells through production of each of the required enzymes in its active form in the cell. Experiments designed to demonstrate cysteine synthesis in intact cells are currently in progress. Initial experiments using the cells of permanent lines (transfected with a similar linked-gene construct) suggest that such cells are indeed capable of synthesis of cysteine if provided with sulphide at a level which is non-toxic to the cells (A.V. Sivaprasad, pers. comm.).

5.5 Coexpression of the cysE and cysM genes in vivo.

Successful coexpression of the cysteine biosynthesis genes of S. typhimurium in the transient and permanent cell culture systems described was a necessary forerunner to attempting the same in vivo. Due to the lack of availability of suitable rumen-specific gene promoters, and a desire to establish an in vivo model of cysteine biosynthesis, a programme of transgenesis using a 6.4 kb SalI / ScaI fragment from the plasmid pLT-EM was begun in mice and sheep. The main aim of the transgenesis experiments described here was to achieve coexpression of the genes in vivo. Transgenesis of sheep with this particular construct is

Figure 5.7 : Quantitation of the cysteine synthesis capacity of cell line pLT-EM.4.

A time course cysteine synthesis assay using 20 µg of protein in an extract from cell line pLT-EM.4 was performed. Time samples were taken at 0, 1, 2 and 5 minutes, then every 5 minutes until 30 minutes had elapsed. The reaction products were fractionated by ascending TLC in the Checchi solvent and the chromatogram autoradiographed as before. The autoradiogram in (a) was analysed densitometrically, the% conversion of serine to cysteine calculated and graphed in (b) as a function of time. In this assay, 1 % conversion of serine to cysteine is equivalent to the conversion of 0.206 nmoles of serine. In the linear range of the assay, the conversion rate is 0.17 % / minute and this equates with 35 nmoles of cysteine / minute/ 20 µg of protein.

a.

NAS–
OAS–
serine–

ysteine–4VP–
origin–

S O N 0 1 2 5 10 15 20 25 30

Time (minutes)
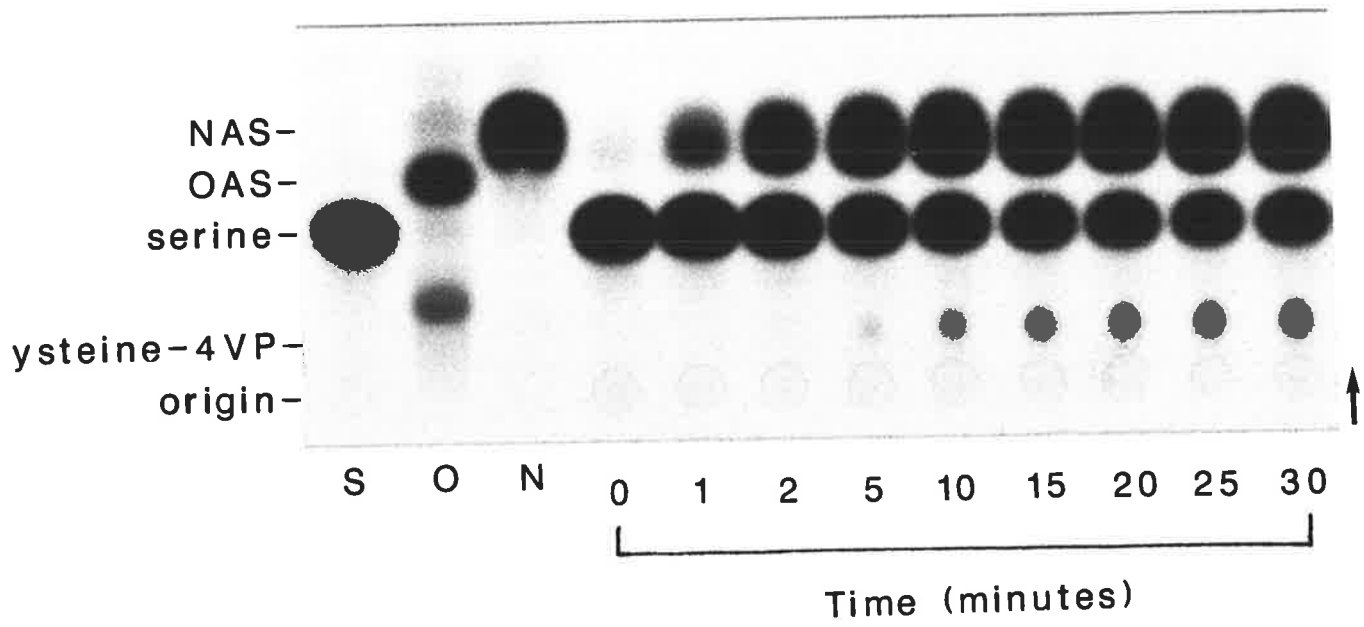
b.

Cysteine Synthesis Time Course : pLT-EM.4
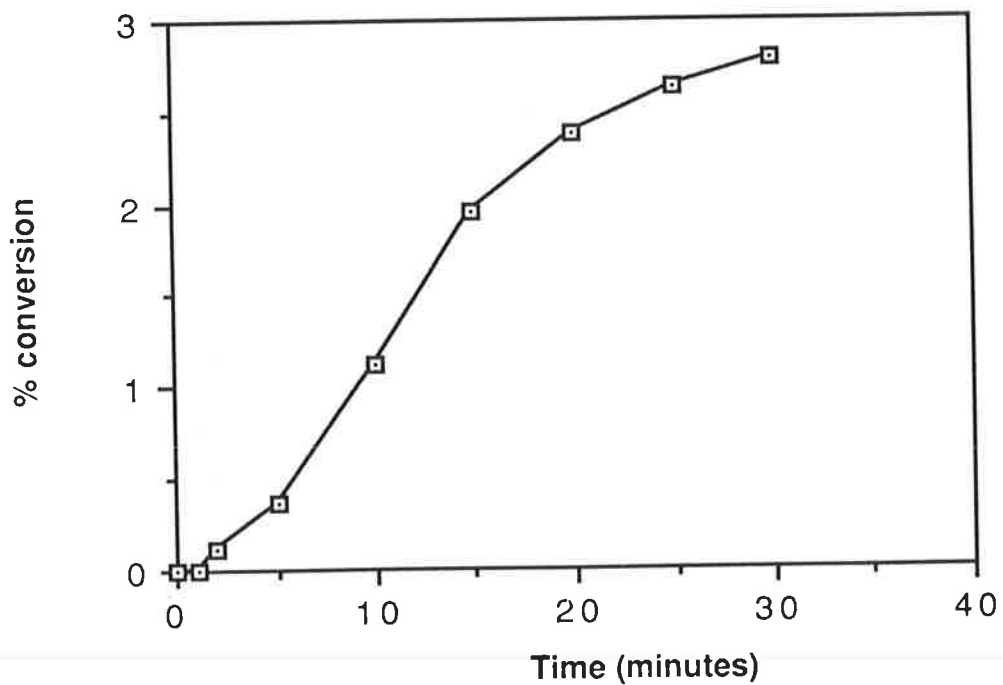
unlikely to provide the desired phenotype (unless the RSVLTR promoter happens to function in the epithelial cells lining the sheep rumen). Moreover, successful coexpression _in vitro_ does not guarantee the same _in vivo_. However, in addition to the RSVLTR-promoted cysE gene expression achieved in a primary sheep foetal lung fibroblast cell line _in vitro_ using the construct pLT-E (see section 4.2 b) data gathered in these transgenesis experiments could give a preliminary indication as to whether the cysteine biosynthesis genes of _S. typhimurium_ can be expressed as efficiently in the sheep itself. The results of transgenesis experiments performed using the linked-gene plasmid construct pLT-EM in mice and sheep are described below.

5.5.1   Coexpression of the cysE and cysM genes in transgenic mice.

Generation of founder transgenic mice with a 6.4 kb SalI / ScaI linked-gene fragment isolated from the clone pLT-EM (see figure 5.2) was achieved as described previously in the case of transgenesis with the cysE gene. In this instance, 13 from 66 of the newborn mice (representing a level of only 19.7 % transgenesis using this larger fragment) were determined to be transgenic by dot-blot hybridization (data not shown). Generally, only 2 to 5 copies of the transgene fragment were integrated; only one mouse, EM-2, had a large number of copies (approximately 100) which were determined to be intact by Southern analysis (data not shown). When a protein extract from the tail tissue of these mice was assayed, none were found to be producing active SAT enzyme in this tissue (data not shown). A positive control extract used in this assay (E.1, from permanent CHO cell line pLT-E.1) exhibited the usual high level of SAT activity.

A comprehensive analysis of protein extracts from 18 tissues of one of the F_0 mice, EM-2, revealed no trace of SAT activity (figure 5.8 a). Analysis of 9 tissues (heart, liver, leg muscle, diaphragm, stomach, sternum, small intestine, foot and skin) from 4 of the other 12 founder transgenic mice also failed to detect SAT activity in any tissue (figure 5.8 b). Protein extracts from only two tissues, the leg muscle (high activity expected) and liver (no activity expected) of the remaining 8 F_0 transgenic mice also showed no SAT activity (data not shown). Again, in each of these analyses, the positive control extract, E.1, displayed

Figure 5.8 : SAT assay analysis of tissue protein extracts from F_0 linked-gene transgenic mice EM-2, EM-18, EM-31, EM-32 and EM-33.

20 µg of protein from 18 tissue extracts of mouse EM-2 (see a) and from 9 tissue extracts of each of the mice EM-18, -31, -32 and -33 (see b) was assayed for SAT activity as described in Methods section 2.2.3. alongside a positive control reaction in which the E.1 extract (from CHO cell line pLT-E.1) was the enzyme source. The reaction products were fractionated by ascending TLC in the Checchi solvent and reference compounds were included in the TLC run. The position of migration of the reference compounds serine (S), OAS (O) and NAS (N), the origin of sample loading and the direction of migration of the solvent are indicated in the figure. Autoradiography of the chromatograms was performed at room temperature for 3 days.

a.

N—
O—
S—

origin—

EM-2

¹⁴C-SER ¹⁴C-OAS ¹⁴C-NAS E.1 Heart Spleen Liver Lung Brain Thymus Eye Leg muscle Kidney Diaphragm Pancreas Stomach Sternum Small Intestine Testis Tail Foot Skin

b.

EM-18   EM-31

N—
O—
S—

origin—

EM-32   EM-33

N—
O—
S—

origin—

¹⁴C-SER ¹⁴C-OAS ¹⁴C-NAS E.1 Heart Liver Leg muscle Diaphragm Stomach Sternum Small Intestine Foot Skin Heart Liver Leg muscle Diaphragm Stomach Sternum Small Intestine Foot Skin

considerable SAT activity. In summary, no SAT enzyme activity was detected in the tissues in which it was expected of any of the transgenic founder mice generated with the 6.4 kb linked-gene fragment. When tail tissue protein from some of the mice was assayed separately for activity of the cysM gene product, OAS sulphydrylase-B, no activity was detected in the extracts either (data not shown). This will be discussed in section 5.6.2.

5.5.2 Coexpression of the cysE and cysM genes in transgenic sheep.

5.5.2.1 Production and identification of transgenic sheep.

Due to the development of an efficient method of assessment of viability of sheep embryos following pronuclear microinjection (including in vitro culture of the embryos prior to transfer to surrogate synchronised ewes; Walker, 1990; Walker et al., 1990) and for the reasons previously given, a sheep transgenesis programme was begun using the same 6.4 kb SalI / ScaI linked-gene fragment as was being used in the mice. Analysis of the lambs produced in this programme commenced at two to three weeks post-partum.

Over a period of 9 weeks of microinjection and embryo transfer, some 230 fertilised sheep ova were processed (Walker et al., 1990). Microinjections were performed with two different amounts of the transgene fragment; approximately 1500 or 3000 copies of the 6.4 kb sequence were delivered to each ovum by pronuclear injection. From this work, 64 lambs were obtained. A proportion of these (5) were stillborn, some others born as twins or triplets (11) died in the first 24 - 48 hours after birth and one was aborted at approximately 120 days gestation. DNA was prepared from the tail of all lambs and analysed by dot-blot hybridization to detect the transgene sequences (figure 5.9 a). As the DNA from the aborted lamb was extensively degraded, the Polymerase Chain Reaction (PCR) was used as an alternative to the dot-blot procedure (data not shown). These analyses suggested that 13 of the 64 were transgenic, approximately 20 % of the F_0 animals. However, of these, 11 were lambs that had died (3 stillborns, 7 that had died soon after birth and one aborted lamb). The transgene copy numbers in the two live sheep, one male (# 181) and one female

5 μg of undigested genomic DNA from each of the F_0 sheep produced in the transgenesis programme was transferred to Genescreen membrane along with the DNA of a copy control series made using plasmid pLT-EM and DNA from a non-transgenic sheep. The membrane-bound DNA was hybridised with the 6.4 kb linked-gene SalI / ScaI fragment isolated from pLT-EM and radiolabelled by Nick-translation with ^{32}P-dATP/dCTP. The filter was washed at 0.2 x SSC / 65 °C and autoradiographed at -80 °C for 16 hours with intensification.

(b) : Southern transfer hybridisation analysis of EcoRI-digested genomic DNA from the F_0 sheep.

10 μg of EcoRI-digested genomic DNA from F_0 sheep identified as possible transgenics by dot blot hybridisation analysis, and from a non-transgenic sheep produced in the programme (83) was fractionated on a 1 % (TAE) agarose gel and transferred to Zetaprobe membrane (vacuum blotted). The filter was hybridised with a 1.0 kb EcoRI / SalI cysE-specific probe (isolated from pGEM-cysE) then washed at 0.2 x SSC / 65 °C and autoradiographed for 48 hours at -80 °C with intensification. The position of the 3.1 kb unit length (cysE) fragment detected in the DNA of sheep 208 is marked (◄) along with those of five fragments detected in the DNA of many of the remaining sheep (◁). A table showing the size (kb) and occurrence of common fragments detected in the DNA of these sheep by the cysE-specific probe is given below.

Sheep no.	84	93	94	96	177	315A	315B	540	807	942A	942B	950	181	208
	-	9.0	9.0	9.0	-	9.0	9.0	-	9.0	-	9.0	9.0	-	(7.8)
Common	-	-	-	5.4	-	-	-	5.4	5.4	-	-	5.4	-	5.4
EcoRI	-	3.8	3.8	3.8	-	-	-	-	3.8	-	-	-	-	3.8
fragments	-	2.7	2.7	2.7	-	-	-	-	2.7	-	-	2.7	-	(3.1)
	-	2.2	2.2	-	-	-	-	-	2.2	-	-	-	-	-

a.

84　93　94　96　177　315A　315B　540　807　942A　942B　950

181　208　　0　0.5　1　2　5　10　25　50　100

Copy no. controls

b.

83　84　93　94　96　177　315A　315B　540　807　942A　942B　950　181　208

kb

23.1-

9.4-

6.6-

4.4-

◀3.1 kb

2.3-
2.0-

0.56-
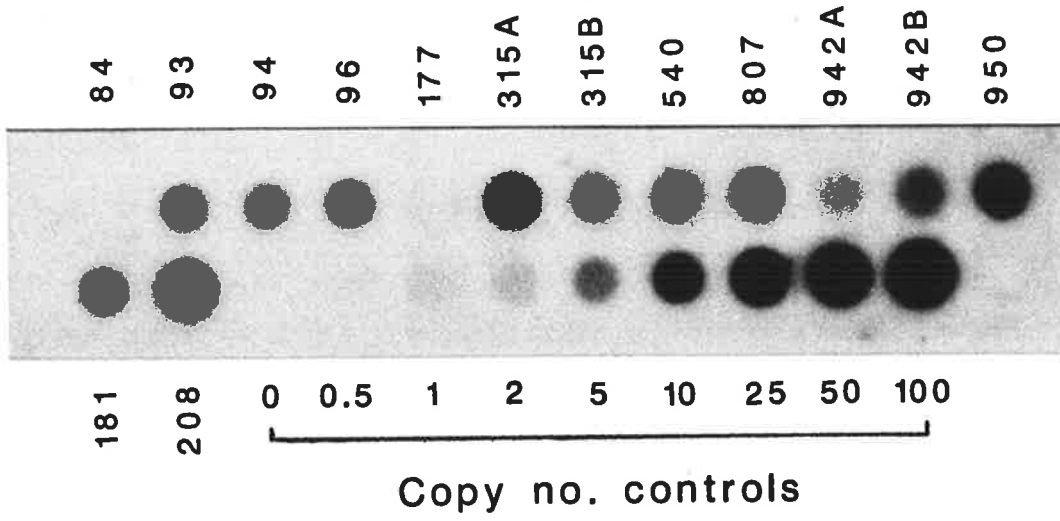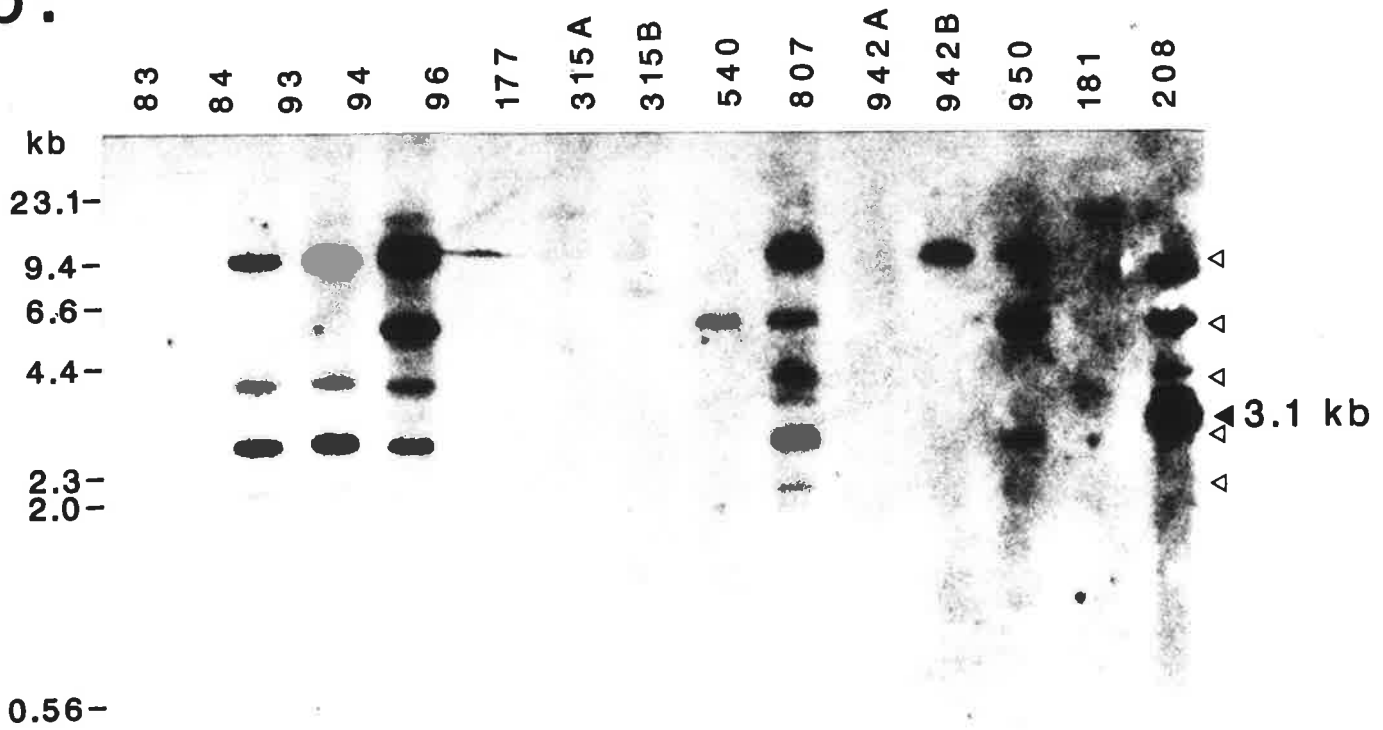
(# 208) were estimated from the dot blot to be approximately 7 and 25 copies respectively.

Following this, the genomic arrangement of the transgene sequences in the sheep was examined by Southern transfer analysis of EcoRI-digested genomic DNA prepared from tail tissue. The probe used was again the 6.4 kb SalI / ScaI fragment. However, whilst this revealed a band pattern in the transferred DNA of many of the sheep, it also produced a high background. Consequently a shorter, more specific detection probe was employed, namely a 1.0 kb EcoRI / SalI cysE-specific gene fragment isolated from pGEM-cysE (see figure 3.2 a) and spanning the entire coding region of the gene. The same Southern filter was stripped of the first probe then reprobed with the cysE-specific probe (figure 5.9 b). Immediately evident in the resulting autoradiogram is the similarity of the band pattern detected in many of the different transgenic sheep. (This was also observed when the 6.4 kb SalI / ScaI probe was used.) The band pattern observed is different to that expected for an EcoRI digest of multiple-copy inserts of the transgene fragment, yet many individual transgenic sheep have the same pattern. Only the live transgenic female 208 has a band pattern which includes the expected 3.1 kb unit-length fragment representing a head to tail array of many intact copies of the 6.4 kb transgene fragment. Both the 3.1 kb cysE and 3.3 kb cysM EcoRI fragments expected to be excised from a tandem array of the 6.4 kb transgene fragment were detected by the 6.4 kb probe used initially (data not shown). In the case of DNA from the other sheep, one or more of 5 common EcoRI fragments are shared by 9 of the sheep. The five common bands are approximately 9.0, 5.4, 3.8, 2.7 and 2.2 kb in size. Five sheep have at least three of these bands. The sizes of the fragments observed suggest that it is unlikely the common band pattern is due to partial digestion of multiple copies of the transgene fragment, whether integrated in a head to head or head to tail array. Other possibilities, including quite specific rearrangement or modification of the transgene sequences either before or after integration, or site-specific integration, are discussed in section 5.6.2. Also evident is a lack of distinct hybridizing species in some DNA samples, namely those from sheep 84, 177, 942 A and 181. A longer exposure of the filter in this case (and when it was probed with other specific fragments) revealed faint bands in the tracks for sheep 84, 177 and 181 but not in the case of sheep 942A. A table

showing the approximate sizes of EcoRI fragments detected by the cysE-specific probe in the genomic Southern of these 12 sheep is given in the legend to figure 5.9 b. As the same amount of DNA was used for each sample, it is possible that the genomes of some of the sheep identified as low copy transgenics by dot blot (i.e. 84 and 177) have multiple loci not detected in this Southern. In the case of sheep 181, 315A and 942A, the disparity between the dot blot and Southern results is assumed to be an artefact due either to the DNA preparation or dot blot procedures.

To further assess the nature of the genomic arrangement observed, particularly in sheep with common EcoRI bands, the Southern filter was stripped and reprobed twice more, first with a hGH 3'end-specific probe (a 487 bp SmaI / DraI fragment; see figure 4.2 b (ii)) then with a RSVLTR 5' end-specific probe (a 444 bp SalI / EcoRI fragment; see figure 4.2 b (ii)). The hGH-specific probe only hybridised to the DNA of sheep 208, detecting five EcoRI fragments, namely the 3.1 kb cysE gene and 3.3 kb cysM gene unit-length fragments and three others (figure 5.10 a). No hybridisation to any of the other sheep DNAs was detected (figure 5.10 a) even after a long autoradiographic exposure (data not shown) suggesting that in these cases the injected 6.4 kb transgene fragment had lost this segment of DNA. However, when the RSVLTR-specific probe was hybridized to the same filter, a band pattern identical to that observed initially with the 1.0 kb cysE probe fragment (and with the 6.4 kb whole probe) was seen (figure 5.10 b). From these data, it appears that the linked-transgene fragment has remained intact in only one transgenic sheep (# 208). It is evident that the RSVLTR sequences have remained associated with the cysE gene (and the cysM gene in the case of sheep 208) during transgenesis. The whereabouts of the cysM gene in DNA of the other transgenic sheep has not yet been determined.

5.5.2.2   Analysis of expression of the transgenes in sheep.

A clearer indication of the function of the transgenes in the sheep was sought by analysis of gene expression in selected tissues of the transgenic ewe 208, as this was the only transgenic sheep in which the construct appeared to have been preserved intact. This animal was sacrificed and tissue samples taken from the heart, liver, lung, brain, leg muscle, kidney, diaphragm, rumen, small

Figure 5.10 : Reprobing of the F₀ sheep Southern filter with hGH-specific and RSVLTR-specific probes.

The Southern transfer filter prepared for the analysis described in figure 5.9 was stripped of the cysE-specific probe and successively reprobed with a 487 bp SmaI / DraI hGH 3'-end probe (see a) and a 444 bp SalI / EcoRI RSVLTR 5'-end probe (see b). Each of the probe fragments was isolated from pLT-E (see figure 4.2 b (ii)) and radiolabelled by Nick-translation with ^{32}P-dATP/dCTP. Autoradiography was carried out at -80 ^{0}C for 2 days (a) and 5 days (b) with intensification. The position of the 3.1 kb (cysE) and 3.3 kb (cysM) unit length fragments detected in the DNA of sheep 208 is indicated in (a) and (b) (◄) and the positions of the five common fragments detected in the DNA of the remaining sheep by the RSVLTR probe are marked on the right side of (b) (◁).

hGH 3'-end probe

a.

3.3 kb
3.1 kb
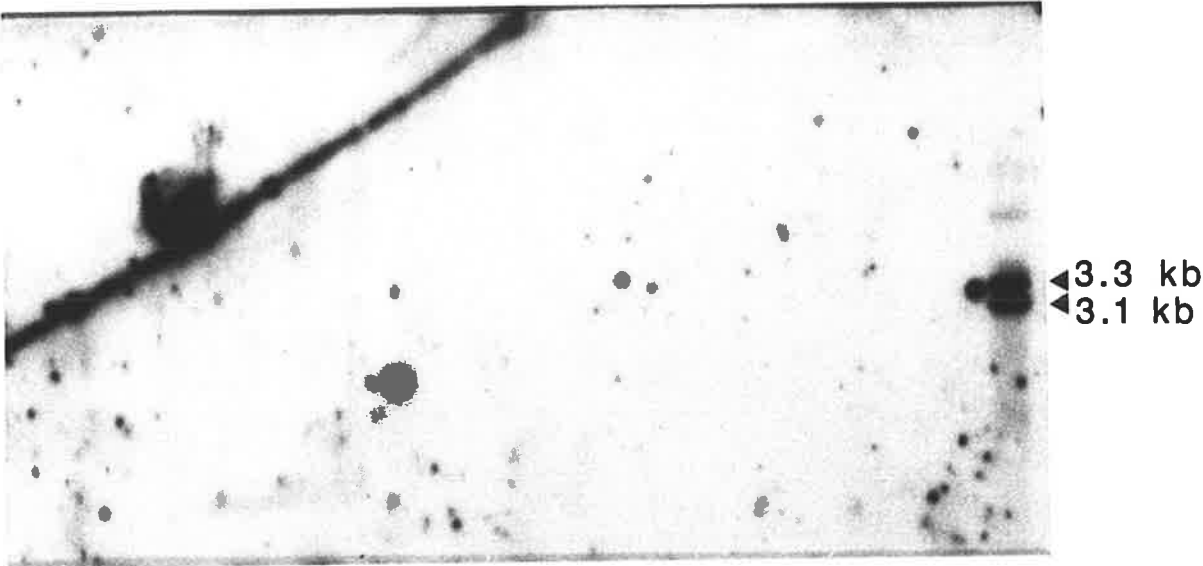
83 84 93 94 96 177 315A 315B 540 807 942A 942B 950 181 208

b.

3.3 kb
3.1 kb
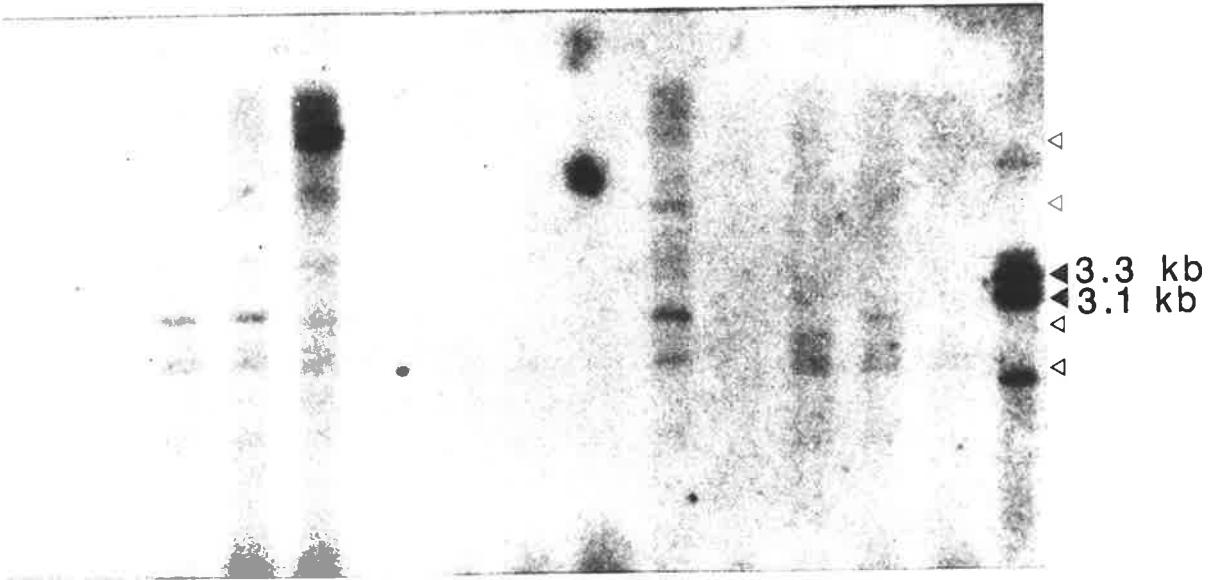
RSVLTR 5'-end probe

intestine and skin, as in the case of analysis of the cysE transgenic mice. In addition, a blood sample was collected. The same tissue samples, including blood, were collected from a non-transgenic ram produced in the transgenesis programme, # 211. These sheep are shown in figure 5.11.

Initially, a crude protein extract was prepared from each of the tissue samples, then 20 μg of total cell protein was assayed for SAT activity (figure 5.12). SAT activity was only detected in the leg muscle and diaphragm of sheep 208, indicating cysE gene expression in those tissues. When analysed separately in a time course assay using 50 μg of protein and twice as much ^{14}C-serine to improve the sensitivity (data not shown) the specific activity of these two protein extracts with respect to the SAT enzyme was determined to be about 1.0 milliunit per milligram of protein. Conversely, SAT activity was not detected in the protein extract made from any other tissue of this sheep, nor in protein extracts made from any of the tissues of the non-transgenic control, 211. A spurious reaction product seen in the lung SAT assay samples from both sheep will be discussed later (see section 5.6.2).

Subsequent to the SAT assay analysis, a cysteine synthesis assay was performed in order to detect OAS sulphydrylase activity (figure 5.13). Only the protein extracts made from lung, leg muscle, diaphragm and rumen of sheep 208 and 211 were analysed, and in this case no supplementary SAT enzyme was added. Assay of the leg muscle, diaphragm and rumen extracts of sheep 211 provided negative controls containing no SAT activity, while assay of the lung extract of each sheep was performed to further test the nature of the enzyme activity detected there. Due to the low level of SAT activity detected in the leg muscle and diaphragm extracts of sheep 208, the time of incubation at 37 °C was increased to two hours in order to improve the sensitivity of the assay. Otherwise, the samples were processed as described previously (see Chapter 2).

In the leg muscle and diaphragm samples of sheep 208, SAT activity was again evident but no cysteine had been synthesized from the OAS produced. As expected, no OAS or cysteine was produced by the leg muscle and diaphragm extracts of sheep 211. Considering the result of this analysis, it is unlikely that a significant level of OAS sulphydrylase activity (if any) is present in the leg muscle or diaphragm extracts of the transgenic sheep 208. The smeared TLC

Figure 5.11 : F_0 transgenic ewe 208 and F_0 non-transgenic ram 211.

The photograph was taken just prior to sacrifice and analysis of tissue extracts from these sheep. Each of the sheep were produced from microinjected ova in the transgenesis programme described, and at the time of analysis were approximately 7 months of age and appeared normal in respect to the South Australian Merino breed phenotype.

211 (Non-transgenic)　　208 (Transgenic)

208

Figure 5.12 : SAT assay analysis of protein extracts derived from tissues of F_0 sheep 208 and 211.

20 µg of protein from 11 tissues (including blood) taken from sheep 208 (top panel) and 211 (bottom panel) was assayed for SAT activity in a 60 minute reaction and the reaction products fractionated by ascending TLC alongside the usual reference compounds and the reaction products of a positive control reaction in which the protein extract from linked-gene CHO cell line pLT-EM.4 was assayed for SAT activity. The chromatograms were autoradiographed for 3 days at room temperature.

(Transgenic)

−NAS
−OAS
−serine
−origin

^{14}C−SER ^{14}C−OAS ^{14}C−NAS EM.4 pRSV.2 Blood Heart Liver Lung Brain Leg muscle Kidney Diaphragm Rumen Small Intestine Skin

−NAS
−OAS
−serine
−origin
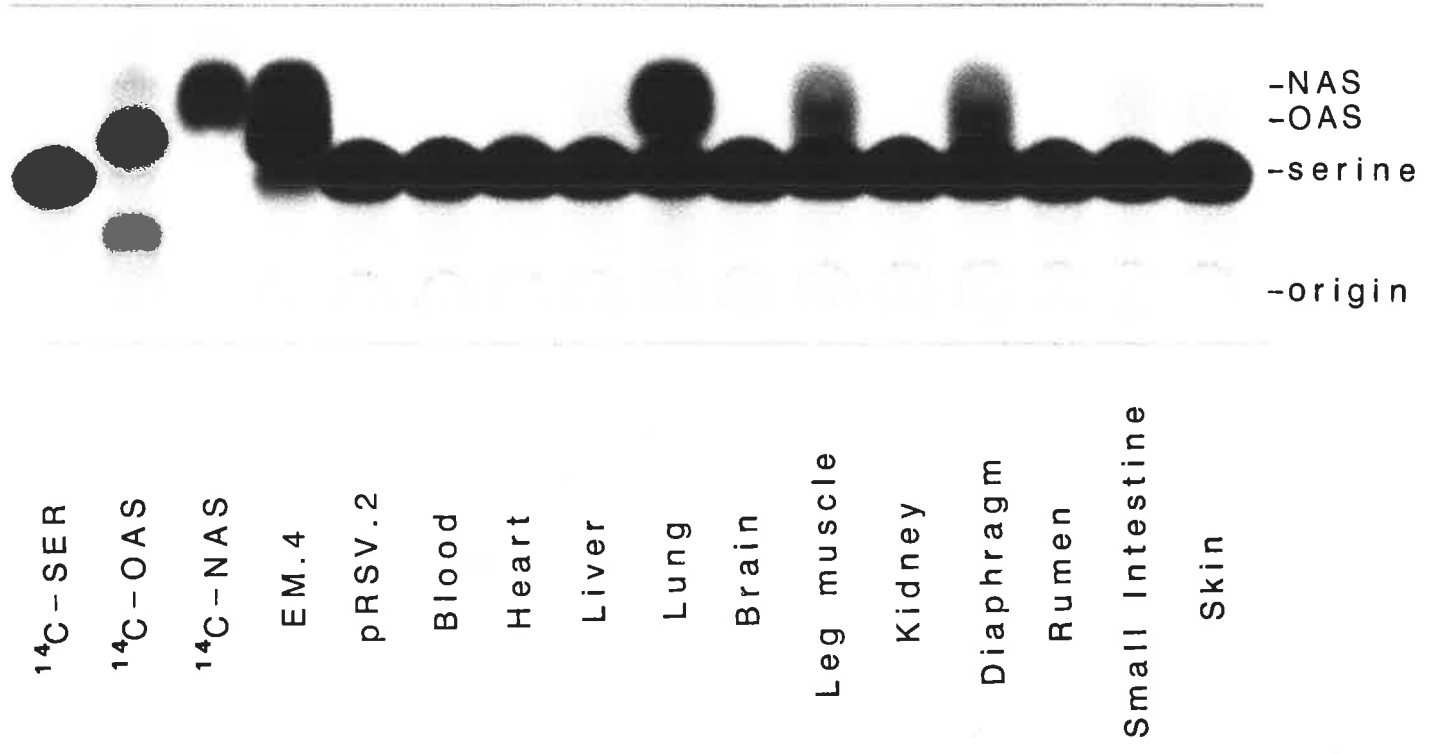
(Non−transgenic)
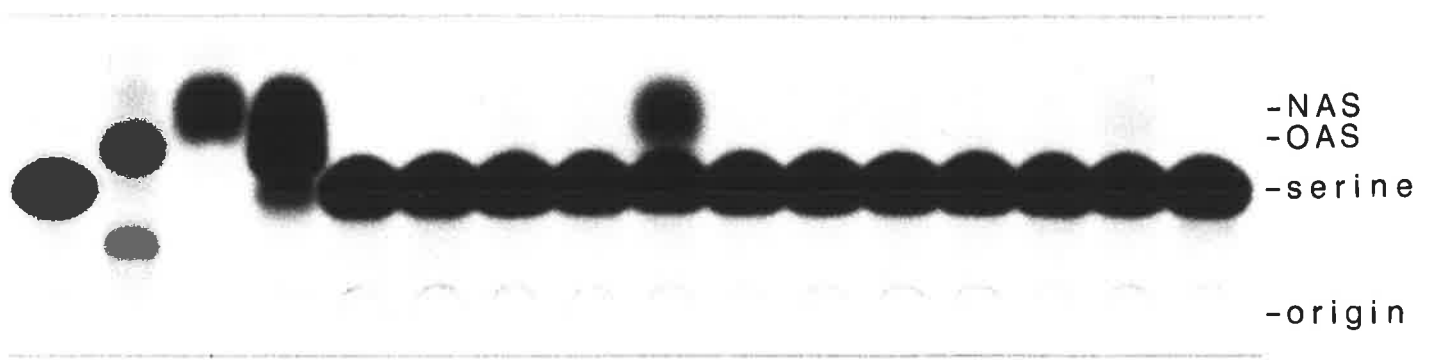
Figure 5.13 : Cysteine synthesis assay analysis of protein extracts derived from tissues of F_0 sheep 208 and 211.

20 µg of protein from each of the lung, leg muscle, diaphragm and rumen tissue extracts from F_0 sheep 208 and 211 was assayed for cysteine synthesis capacity in a 2 hour reaction. The reaction products were fractionated by ascending TLC as described previously and the chromatogram autoradiographed for three days at room temperature. The positive and negative control samples (cysteine synthesis assays of the extracts from cell lines pLT-EM.4 (EM.4) and pSRV.2) were co-fractionated with the test samples and reference compounds.
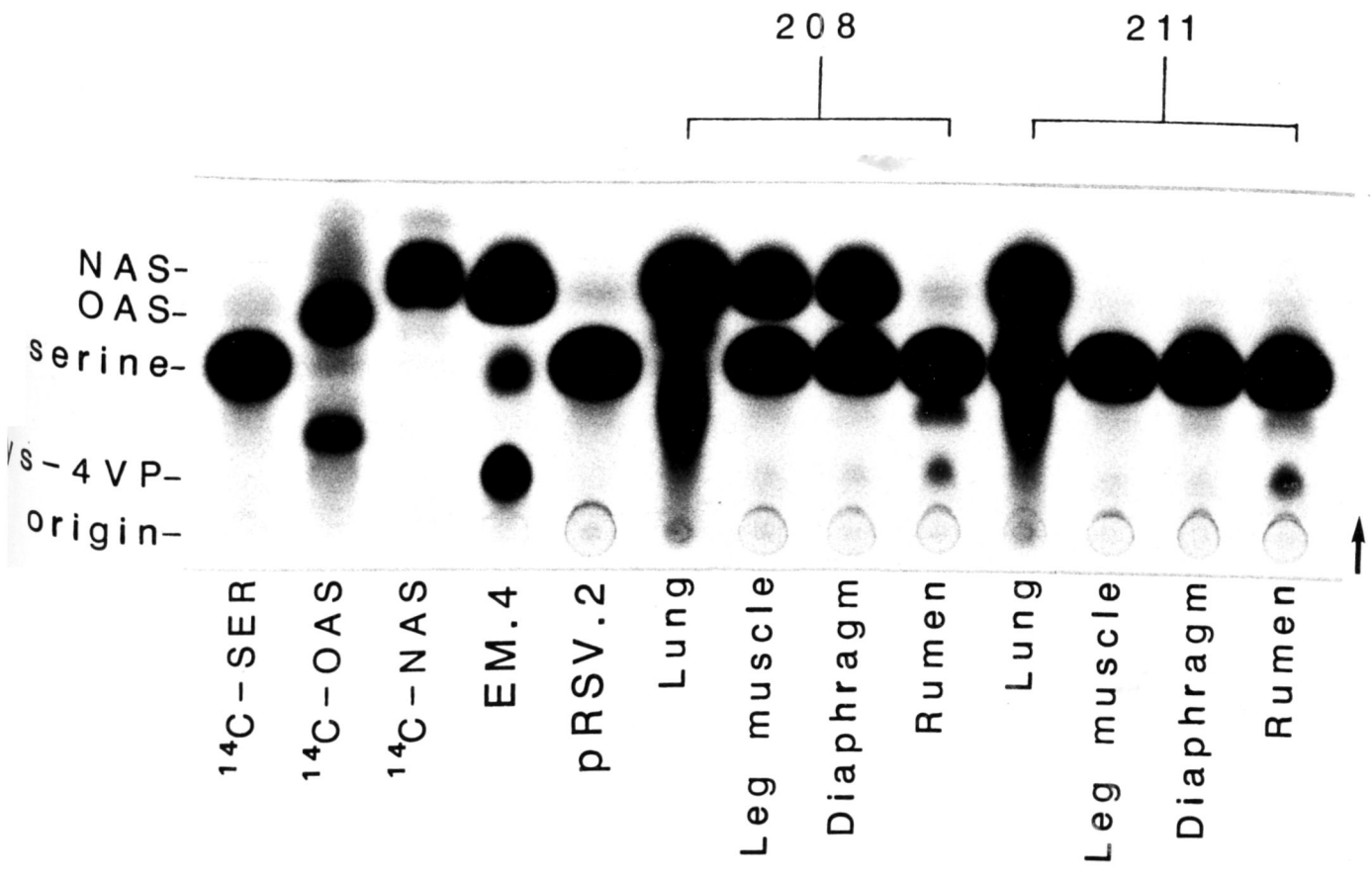
fractionation pattern produced in the case of the assays of the lung extracts of sheep 208 and 211 are difficult to interpret, but again, the supposed OAS / NAS spot has migrated further than expected (see Discussion section 5.6.2). In addition, each of the rumen assay samples features a spot which co-migrates with the cysteine-4VP spot present in the positive control sample (an assay of the protein extract from linked-gene cell line EM.4). In the absence of detectable SAT activity in these extracts (see figure 5.12) it would seem that this product is unlikely to be cysteine synthesized as a result of transgene expression (see Discussion section 5.6.2).

In order to confirm the gene expression detected by enzyme assay, and to resolve the spurious enzyme activities detected in the lung and rumen tissue samples of both the transgenic and non-transgenic sheep, detection of cysE and cysM transcripts was undertaken via RNA protection analysis (figure 5.14). Each of the probes used was capable of detecting both cysE and cysM transcripts, by virtue of the 40 base RSVLTR region common to the 5' end of each transcript in addition to larger, distinctive fragments from their respective genes. A low level of cysE gene transcripts were detected in the leg muscle and diaphragm RNA samples from transgenic sheep 208, but not in the lung or rumen RNA samples (figure 5.14). This is evident from the 280 base protected RNA species present when the cysE pLTRE-B probe was used (figure 5.14 a) and the faint 40 base band when the cysM pLTRM-B probe was used (figure 5.14 b). By comparison with the positive controls used and by densitometry, cysE transcript levels in these tissues were estimated to be approximately 1.6 and 3.1 pg per 10 µg of total cellular RNA, respectively. No cysE transcripts are evident in any of the RNA samples of the non-transgenic sheep, 211 (figure 5.14 a and b).

In contrast, neither the pLTRE-B nor pLTRM-B antisense probes detected cysM gene transcripts in any of the RNA samples, either from sheep 208 or sheep 211. One would expect 40 base (figure 5.14 a) or 250 base (figure 5.14 b) protected fragments to be detected if using these respective probes. Although not at the absolute limit of sensitivity (i.e. polyadenylated RNA could be isolated from these samples and analysed in this fashion) this finding is consistent with the result of the cysteine synthesis assays, in which no OAS sulphydrylase activity was detected (see figure 5.13). At best, if any OAS sulphydrylase

Figure 5.14 (a) : RNA protection analysis of cysE gene expression in total RNA prepared from tissues of F_0 sheep 208 and 211.

10 µg of RNA isolated from the lung, leg muscle, diaphragm and rumen tissue samples taken from F_0 sheep 208 and 211 was analysed by RNA protection for the presence of cysE gene transcripts using the pLTRE-B antisense probe which had been radiolabelled with ^{32}P-rUTP. The products of the protection reactions were fractionated in a denaturing 6 % (19:1) Urea-PAGE system and the gel autoradiographed for 48 hours at -80 ^0C with intensification. The probe controls Y+ and Y- (described in Methods section 2.2 2) and positive controls (50 and 500 pg of non-radiolabelled cysE sense transcript diluted into 10 µg of yeast tRNA) were co-fractionated with test samples on the gel; only one tenth of the Y+ and Y- controls was loaded onto the gel. The position of 285 base protected cysE RNA detected in the leg muscle and diaphragm samples is indicated on the right side of the figure. No cysM transcripts (expected to be about 40 bases in size if protected by this probe) were detected in this analysis.

a.
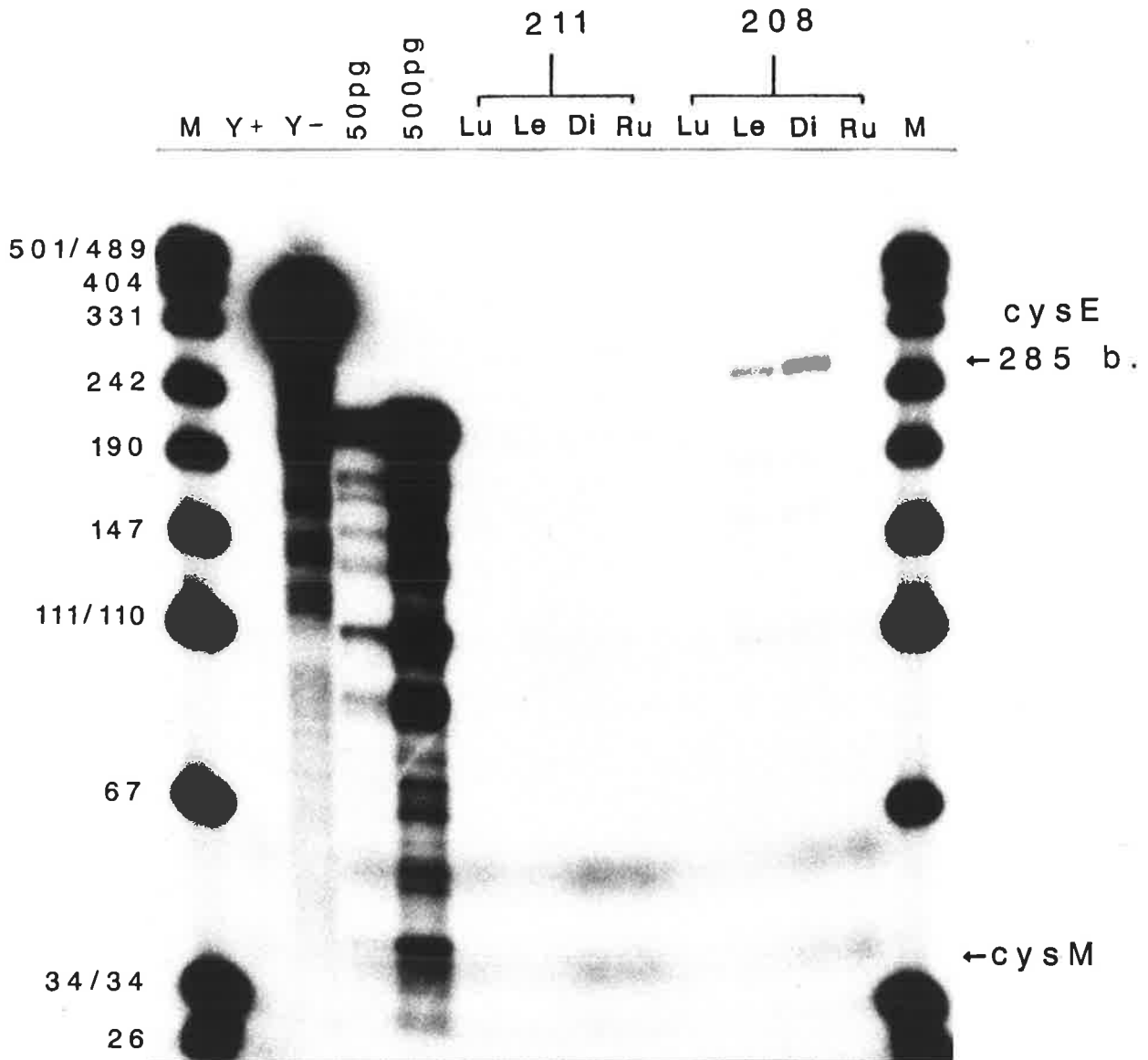
Figure 5.14 (b) : RNA protection analysis of cysM gene expression in total RNA prepared from tissues of F_0 sheep 208 and 211.

Ten μg of RNA isolated from the lung, leg muscle, diaphragm and rumen tissue samples taken from F_0 sheep 208 and 211 was analysed by RNA protection for the presence of cysM gene transcripts using the pLTRM-B antisense probe which had been radiolabelled with ^{32}P-rUTP. The products of the protection reactions were fractionated in a denaturing 6 % (19:1) Urea-PAGE system and the gel autoradiographed for 48 hours at -80 ^{0}C with intensification. The probe controls Y+ and Y- (described in Methods section 2.2 2) and positive controls (500 and 50 pg of non-radiolabelled cysM sense transcript diluted into 10 μg of yeast tRNA) were co-fractionated with test samples on the gel; only one tenth of the Y+ and Y- controls was loaded onto the gel. The position of 40 base protected cysE RNA detected in the leg muscle and diaphragm samples is indicated on the right side of the figure. No cysM transcripts (expected to be approximately 250 bases in size if protected by this probe) were detected in this analysis.
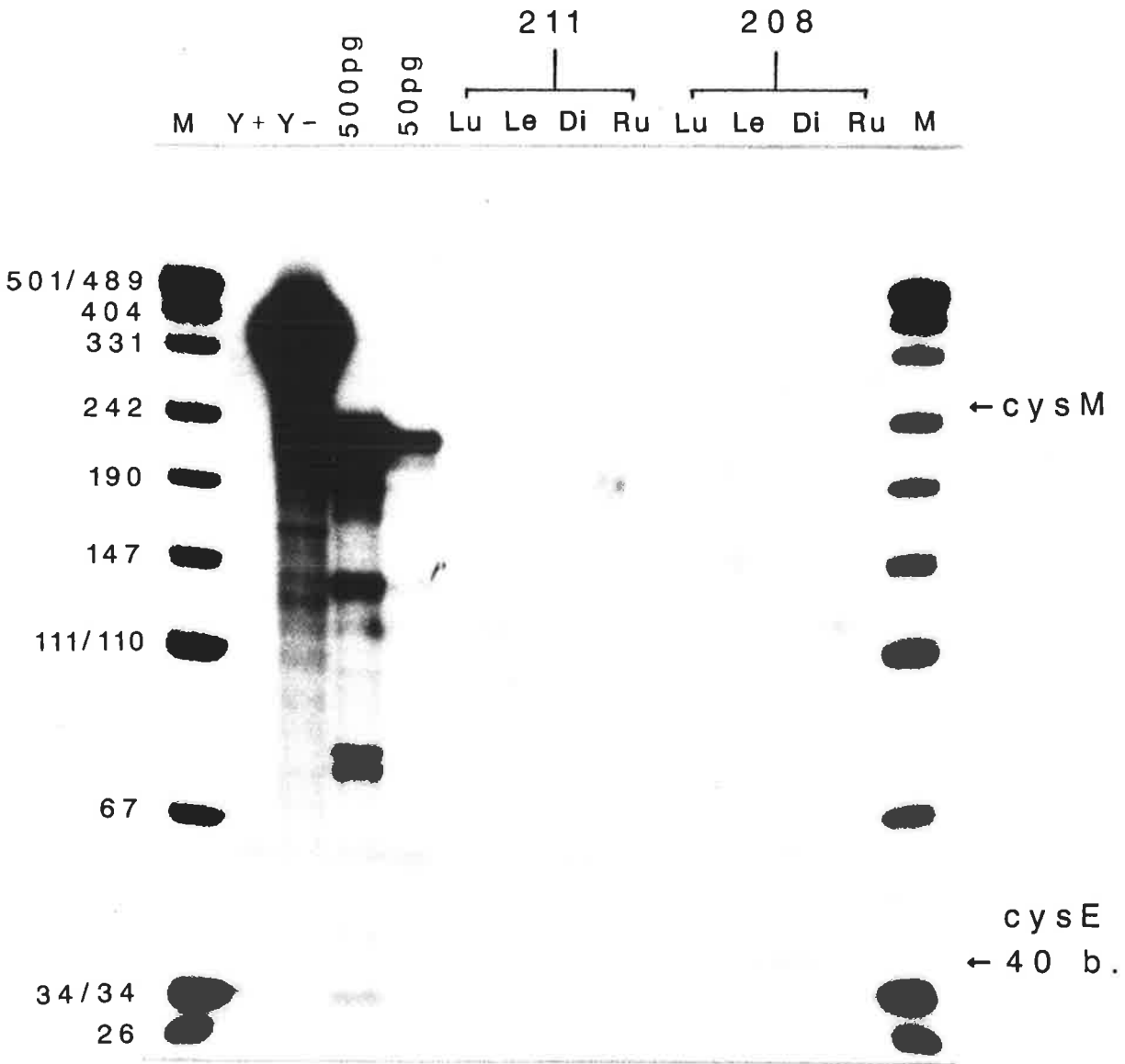
b.

monomers are produced, they are insufficient in number to allow detectable enzyme activity in the assay system employed.

Considering the likelihood of tissue degradation in the case of the stillborn lambs, and others which had died whilst not in attendance, analysis of expression of the transgenes in these sheep was only attempted by assay of SAT activity from a protein extract from tail tissue. No SAT activity could be detected in the tail tissue. When assayed separately for OAS sulphydrylase activity, by addition of E.1 extract to the assay mixtures to provide the substrate OAS enzymatically, no activity was detected (data not shown).

Discussion.

5.6 Coexpression of the cysteine biosynthesis genes of Salmonella typhimurium in mammalian systems in vitro and in vivo.

If the results of coexpression of the cysteine biosynthesis genes in cell culture experiments and preliminary data gathered from the first sheep transgenesis experiment are considered in relation to the primary aim of the thesis, i.e.to establish a new de novo cysteine biosynthesis pathway in the sheep rumen epithelium, then it is quite feasible that this aim may be achieved. Improvements are obviously neccessary to optimise the specificity and level of gene expression attained in vivo in the sheep, yet the level of expression of active enzymes measured in the cell culture experiments clearly indicates the potential of the genes in use.

5.6.1 Coexpression of the cysE and cysM genes in vitro.

Initially in transient expression experiments, then via the creation of permanent cell lines using the plasmid construction pLT-EM, coexpression of the cysteine biosynthesis genes of Salmonella typhimurium was achieved. The level of gene expression measured in the case of one of the permanent cell lines, EM.4, was quite high; 20 μg of a protein extract from this cell line supported a rate of cysteine synthesis of about 35 nmoles per minute (i.e. 50.4 μmoles of cysteine per day). In order to produce a minimum of 1.5 grams (12.4 mmoles) of cysteine

per day (see section 1.3) only 4.9 milligrams of protein of this specific activity would be required. Considering the size of the sheep rumen and the extent of the convoluted epithelial lining of that organ, it could reasonably be expected that a similar level of gene expression to that found in line EM.4 would easily support cysteine synthesis of this magnitude in the sheep.

The finding of different levels of expression (at the level of transcription) between the tandem cysE and cysM genes in the few permanent cell lines studied here implies the need for improved control of expression of the genes in linked-gene constructs. This could be achieved by the use of more determinate regulatory DNA elements and these will be discussed further in Chapter 6. A disparity in the level of expression of the cysE and cysM genes has also been observed in the case of two new series of linked-gene permanent cell lines made in this laboratory (A.V. Sivaprasad, unpublished data; Jenkins, 1990). Each of these new constructs is different to pLT-EM in that the cysM gene is cloned in tandem but positioned 5' to the cysE gene, the idea being to determine whether gene arrangement in the linked gene constructs has a bearing on the expression levels achieved for the individual genes. In addition, these two constructs differ from each other with respect to the 3' processing sequences that they contain. One features each coding region flanked on its 3' end by the hGH sequences as found in the original construct pLT-EM. The other has the upstream mouse sequence (UMS; a polyadenylation and termination sequence; Wood et al., 1984; Heard et al., 1987) placed 3' to each gene. In this instance, the need to minimise possible transcriptional interference in the gene pair (more critical when the linked-gene construct is used in transgenesis) was being investigated.

In comparison with the pLT-EM cell lines, initial data regarding the level of expression of the genes of these new constructs in permanent CHO cell cultures suggests that neither alteration of the gene arrangement (with cysM now preceding cysE) nor the use of reportedly more efficient transcription termination sequences (i.e. the UMS sequence) necessarily resolves the disparity in the level of expression of the genes in the pair in this system. However, as in the case of the pLT-EM cell lines, only a small number of cell lines derived from each of these new linked-gene plasmids have been analysed. Consequently, the nature of the inequality of expression of these linked genes remains to be more fully

examined. It is still possible that the effects seen in the pLT-EM cell lines (and those of the two new series) where the number of copies of integrated plasmids are low (2 to 5 copies in most cases) are due largely to the influence of the host genome in the locale of the integration site. In some cases it is true that the cell lines which express the gene pair well are those with the highest number of plasmids integrated (e.g. line EM.4, and other cell lines of the new series). Although logical, this trend is questionable in the absence of more data, and cell lines with a higher number of integrated plasmid copies certainly do not exhibit copy number-dependent expresssion (e.g. the level of cysteine synthesis in line EM.4 is in fact far higher than could be expected according to the fold increase in integrated plasmid copy number).

As an adjunct to this, manipulation of expression of the genes to achieve specific levels of enzyme activity may be necessary to facilitate maximal cysteine synthesis. In particular, it is clear that the level of SAT enzyme activity in relation to OAS sulphydrylase activity in a cell is critical. It was noted that a cell line with a low level of SAT enzyme activity and a moderate level of OAS sulphydrylase activity (i.e. cell line EM.6) was capable of significant levels of cysteine synthesis. Where the possibility of loss of some OAS (by acyl migration to NAS at neutral pH) exists in vivo, it seems sensible to maintain the specific activity of the second-step enzyme, OAS sulphydrylase, at a higher level than the first-step enzyme, SAT, in the cell. This may be achieved by limiting the expression of the cysE gene through the use of a less efficient promoter of transcription, or conversely to enhance the expression of the cysM gene. Although expression of high levels of the cysE gene product in transgenic mice suggests that this should not present a problem to cell metabolism, another use of such a limitation would be to diminish the possibility of intracellular toxicity caused by accumulation of OAS or NAS.

Another factor which has a bearing on the level of SAT activity in the cell is the feedback inhibition of this enzyme by quite low levels (1.1 uM; Kredich and Tomkins, 1966) of L-cysteine. In vivo, this self-regulation may be an advantage if overproduction of cysteine proves to be toxic to the cell. However, if this is not the case and it is desirable to produce cysteine at higher levels, then the use of a cysE gene from a mutant Escherichia coli strain in which the enzyme is

about 10-fold less sensitive to feedback inhibition by L-cysteine (Denk and Bock, 1987) may be pursued.

To summarise the in vitro expression of the cysteine biosynthesis genes reported here, first with respect to the cysE gene, the relationship between RNA and enzyme levels established in the permanent cell lines suggests that this gene is at no disadavantage in the cell with respect to (i) transcript stability or (ii) catalytic activity of the protein product. This is yet to be established in the case of the cysM gene. Poor translation efficiency was noted earlier for cysM transcripts in the RRL but in the more recent constructs containing the cysM gene, this has been improved by optimisation of the initiation context in the 5'-untranslated region of the cysM gene (A.V. Sivaprasad, unpublished data). In addition, the OAS sulphydrylase-B enzyme has a lower affinity for its substrates (i.e. K_m values of 7.0 mM and 3.6mM for OAS and sulphide respectively; Kredich, 1987) compared to the SAT enzyme (i.e. K_m values of 0.56 mM and 0.11 mM for serine and acetyl CoA respectively; Kredich and Tomkins, 1966). Other than these considerations, the in vitro coexpression results presented here suggest that the cysM gene product is stable and active in a mammalian environment. As expected, in the case of each gene in each cell line, the level of transcript detected by RNA protection was generally mirrored by the level of enzyme activity of that gene product, and this correlation should be achievable in vivo.

5.6.2 Coexpression of the cysE and cysM genes in vivo.

Following the initial finding that the linked-gene construct was successful in terms of coexpression of the cysE and cysM genes in vitro, it was used in transgenesis of mice and sheep. It was hoped that analysis of the behaviour of the linked transgenes contained in pLT-EM would demonstrate that the genes could be coexpressed in vivo and particularly, that they were suited to use in the sheep. Until the work reported here was undertaken, expression of these genes in the sheep remained to be tested.

Before undertaking the sheep transgenesis programme, it was thought that mouse transgenesis with this construct would rapidly provide data regarding coexpression of the genes in vivo, although it was recognised that the results may not necessarily be applicable to sheep. In fact in this case, and for reasons yet to

be determined, none of the transgenic founder mice produced detectable levels of SAT or OAS sulphydrylase activity in any of the tissues examined. Analysis of transcription of the cysE and cysM genes in specific tissues (i.e. mesodermal tissues) of these transgenic mice remains to be carried out. However it is assumed that if transcription occurred, it was at a level too low to permit the production of any active enzyme molecules. This is inferred from the results of the enzyme assays and RNA analyses of the linked-gene permanent cell lines (and sheep) where transcript levels as low as about 1.0 pg of transcripts per 10 µg of total cellular RNA are still sufficient to provide easily detectable enzyme activity. Considering the number of founder animals produced, and that the RSVLTR promoter is known to be nonfunctional in the developing mouse embryo (Overbeek et al., 1986) lethality due to expression of the cysteine biosynthesis genes during embryogenesis was ruled out as a reason for the lack of expression observed.

Concomitant with the mouse transgenesis experiments, a sheep transgenesis programme was begun to establish expression of the linked transgenes in the sheep. It was recognised that the RSVLTR gene promoter used in the plasmid pLT-EM is not the ultimate choice for sheep transgenesis, considering the requirement for expression of these genes in the sheep rumen (the site of availablity of the substrate, sulphide) and that only trace amounts of activity of the cysE gene product were measured in similar tissues (the stomach and small intestine) of the cysE transgenic mice (see figure 4.6). However, in the absence of a proven rumen-specific promoter, the RSVLTR promoter, known for its high level of activity in specific tissues, was used in an attempt to direct the expression of the linked transgenes in sheep. As in the case of the mouse transgenesis, a reasonable proportion of the lambs produced in the sheep transgenesis programme described here were identified as carrying the transgene fragment. In a Southern analysis, the EcoRI-digested DNA of the live transgenic female sheep, 208, gave a pattern of bands expected for a multi-copy tandem integration of the 6.4 k.b. fragment. In contrast, this analysis also revealed that in many of the remaining lamb genomes, the 6.4 k.b. SalI / ScaI fragment had sufferred specific rearrangements. These resulted in a common pattern of EcoRI restriction fragments of unexpected size in many of the genomes. The band

patterns observed were inconsistent with those expected from partial digests of the 6.4 kb linked-transgene fragment, whether in a head to head or head to tail array. It is possibly significant that these particular lambs were amongst those (11 from 17) that died at or soon after birth. However, the reason for their death or for the DNA rearrangement observed may only be surmised, and the rearrangement may be related either to the processes involved with integration of the transgenes, or to processes which occurred in the cell nucleus at the time of death of the lambs.

In a sense, the former of these possibilities is favoured, for one feature noted in the DNA rearrangement was the loss of the hGH 3' end sequences from the integrated DNA detected in the genomes of all sheep except that of the live transgenic 208. The RSVLTR sequence was shown to have remained associated with both the cysE and cysM genes in transgenic 208 and with at least the cysE gene in the rearranged DNA of the other sheep. DNA rearrangement related to growth hormone sequences has been observed previously and found to be the result of interaction between repetitive DNA elements in the 3'-flanking DNA of the endogenous growth hormone gene and similar elements contained within the growth hormone transgene fragment in use (Vize, 1987). In this regard, it is interesting that the sheep growth hormone (sGH) gene, like other growth hormone genes, is flanked by middle-repetitive DNA. At the 3' end of the sGH gene, two elements belonging to the C-A3 family of repeated DNA, thought to be a form of transposable element or retroposon (Rogers, 1985) are found (Byrne et al., 1987). Although the 3'-flanking region within the hGH sequence used here in the linked-transgene fragment is not very large, small regions of homology between this and the 3'-flanking DNA of the sGH gene are evident. In fact six 9 bp DNA sequence motifs sharing homology to the 9 bp direct repeats which flank the C-A3 elements (found in the sGH gene 3'-flanking DNA, and with the sequence TGGGAAATG) exist in this region. The consensus for these is TGGG(A/T)(C/T)NAC, hence all share 4/9 bp homology to the sheep consensus (two of the six elements share 5/9 bp homology, one has 6/9 and another 7/9 bp homology). In addition, a sequence resembling an RNA polymerase III promoter element, again as seen in the C-A3 elements flanking the sGH gene, is present here. If the transgene was involved in specific DNA rearrangements at the sGH

gene locus through the interaction of these hGH 9 bp elements with the sGH 9 bp direct repeats, or by C-A3-type retroposon activity, then the function (if any) of this gene in the developing lamb could be compromised. Interactions such as these could be tested directly by using an sGH gene-specific probe to see if the transgene sequences in those particular lamb genomes are associated with the endogenous sGH sequences or have been transposed to other specific sites.

Although it is mere conjecture to suggest a role for the transgene sequences in disrupting the structure and function of the endogenous sGH gene, and ultimately in causing the death of these lambs, this possibility cannot be ruled out. Unfortunately, while placentally derived growth hormone is known to function in the foetus, the time of commencement of expression and the precise functional role of the endogenous sGH gene in the developing lamb is unknown. At least, the results do suggest disruption of a vital function by the transgene sequences. The presence of a sequence (not necessarily the sGH gene) within the genome with which the transgene can preferentially interact would also help to explain the high proportion of transgenic lambs produced in this experiment and their premature deaths.

One factor not yet considered here is the possible role of the RSVLTR sequence in mediating the DNA rearrangement. In this instance, it would seem likely that a rearrangement event could occur prior to integration of the transgene into the genome, and in the manner discussed in Chapter 4 (see Discussion section 4.4). As yet, insufficient data has been gathered to support or negate this possibility. However, given the difficulty in obtaining transgenic mice expressing the genes in the linked-gene construct, and the DNA rearrangements observed in the case of all but one of the transgenic sheep produced, the interactions between these sequences (i.e. the RSVLTR and hGH sequences) and the host genomes into which they are placed should be carefully monitored in transgenesis experiments where they are used. We plan to use alternative tissue-specific gene promoters and 3' end processing sequences to replace these as discussed in Chapter 6.)

Despite the difficulties encountered in achieving in vivo coexpression of the cysteine biosynthesis genes, it is encouraging that of only two live transgenic sheep produced in this experiment, the female 208 displayed tissue-specific

expression of at least one of the two genes. Analysis of gene expression in the live transgenic male sheep 181 will be carried out when the genomic arrangement of the transgene fragment is more accurately determined. Although the expression of an enzymically active product from the cysE gene is only at a low level, and not in the desired tissue (i.e. the rumen) this observation suggests that the the intracellular processing machinery found in the cells of sheep tissues in vivo can support the synthesis of SAT enzyme monomers, and that these are then free to aggregate to form the enzymically active polymeric species. Activity of the bacterial cysE gene product had previously been measured in a sheep primary cell line (see Chapter 3) but demonstration of this activity in vivo in sheep was of prime importance because a non-functional first-step enzyme would negate the aim of this work.

As stated previously, the pattern of gene expression observed in the tissues of sheep 208 was largely as observed in the cysE transgenic mice. In a SAT assay analysis, expression of the cysE gene was detected at a low level in the leg muscle and diaphragm, but not in any of the remaining eight tissues (or the blood sample) taken from the transgenic sheep 208. In a situation such as this, in which the gene expression is very poor in the skeletal muscle (which is expected to exhibit the highest activity of all of the mesodermally-derived tissues; Overbeek et al., 1986) it is consistent that expression in other mesodermal tissues of the animal is extremely low and possibly undetectable. In relation to the low level of SAT activity measured in protein extracts derived from the leg muscle and diaphragm of sheep 208 (of the order of 1.0 milliunit per milligram of protein) a correspondingly low level of cysE gene transcripts was detected in RNA isolated from these tissues. This correspondence, as determined for cysE gene expression in the linked-gene permanent cell lines, also indicates that the SAT molecules produced are relatively stable in sheep cells in vivo.

The lack of expression of the cysM gene in sheep 208, first assumed due to the inability of the leg muscle and diaphragm extracts to support cysteine synthesis under standard assay conditions, was confirmed by the results of RNA protection analyses, in which no cysM gene transcripts were detected in RNA isolated from these tissues.

The importance of the enzyme assays in assessing gene expression in these experiments was considerable. Due to this, and for the sake of future experimentation, the results of the assays of the lung and rumen tissue samples require explanation. In the SAT assays performed, one feature common to the lung protein extract of both sheep 208 and 211 was conversion of the ^{14}C-serine to an unknown substance which migrated a little more rapidly in the TLC than the expected OAS / NAS spot and produced a more diffuse spot in the TLC (see figure 5.12). It is possible that some endogenous enzyme activity capable of such a conversion exists in the lung of the sheep, although there is no such activity in any of the other tissues assayed in this way. SAT assay results obtained from the lung tissue extracts of the cysE transgenic mice (see figure 4.6) suggest that in the case of transgenic sheep 208, transgene expression in the connective tissue of the lung may give rise to detectable SAT activity. However, this would not account for the substrate conversion seen in the assay of the lung protein extract from the non-transgenic sheep 211. Alternatively, as these animals grazed together, it is quite likely that the substrate conversion seen was due to the action of the enzymes of common parasites present in the lung tissue of each animal. If this was the case, and the product was OAS (and resolved a little differently in the TLC due to the presence of substances present in the lung tissue, for example the surfactant produced there) this could be distinguished from authentic transgene expression and SAT activity by examination of cysE gene expression at the level of transcription (see below).

Similarly, the apparent cysteine synthesis observed in assays of the rumen samples from sheep 208 and 211 must be explained. Again, whilst it was possible that an endogenous rumen enzyme was acting on serine, it is more likely that the enzymes of parasites (in this case adherent bacteria which are known to inhabit the wall of the rumen and are expected to be present in small quantities in a crude protein extract made from this tissue) were responsible for the small amount of conversion seen. In any case, a result such as this could cause a complication, particularly considering that a determination of the cysteine synthesis capacity of cells of the rumen epithelium will be sought in future experiments.(N.B. Some microorganisms possess enzymes which are capable of synthesis of cysteine directly from serine in a single step (Schlossmann and Lynen, 1957). Contrary to

expectation however, the gene for an enzyme such as this does not form the main subject of this work because the equilibrium of the particular reaction is in favour of synthesis of serine from cysteine (Huovinen and Gustafsson, 1967).)

In order to resolve each of these spurious results, RNA protection analyses were employed (see figure 5.14 a and b). In the absence of detection of any cysE or cysM transcripts in the RNA samples from the lung or rumen tissue of either animal, the enzymatic conversion of serine in the SAT and cysteine synthesis assays to products resembling those expected in the respective assays, are deemed artefacts peculiar to these tissue samples and the assay systems employed here. These should be noted carefully in regard to their possible interference in the interpretation of results of future assays of crude protein extracts, especially where rumen tissue is to be analysed as the preferred site of expression of the cysteine biosynthesis genes.

Finally, from the limited amount of data obtained from sheep 208, it is impossible to ascertain whether the RSVLTR promoter (which by all accounts is one of the more powerful eukaryotic promoters available) has been hindered in any specific way other than by genomic location alone. Although disappointing, the lack of expression of the cysM gene, evidently intact according to the Southern analysis, may also be the result of hindrance by the genomic DNA immediately flanking the integration site. This is likely considering that differential expression of the linked genes has been observed previously in vitro in the pLT-EM permanent cell lines (e.g. cell line pLT-EM.3, where the cysE gene is expressed well, but expression of the cysM gene is barely detectable, or cell line pLT-EM.6 in which the converse is true). It is hoped that analysis of the expression of the transgenes in sheep 181 and other transgenics will yield further information regarding in vivo coexpression of the cysteine biosynthesis genes in the linked-gene construct, pLT-EM (specifically whether lack of cysM gene expression is a general phenomenon or characteristic to this construct) and in turn, information which is useful in achievement of the ultimate aim of cysteine biosynthesis in the sheep.

Chapter 6

General Discussion.

Chapter 6 : General Discussion.

6.1 Introduction.

Many achievements described in this thesis indicate the possibility of reaching the goal of improved wool growth in sheep. These include (i) tissue-specific expression of high levels of the cysE gene product _in vivo_ in transgenic mice, (ii) expression of the cysE and cysM genes, whose products facilitate the two-step biosynthesis of cysteine _de novo_, in a mammalian cell culture system and (iii) tissue-specific expression of the cysE gene in transgenic sheep. Whilst substrate availability would limit the function of cysteine biosynthesis gene products expressed in these situations, the consequences of the production of larger quantities of cysteine intracellularly are unknown. Indeed, the more fundamental questions of the effect of higher levels of intracellular cysteine on the metabolism of the cell, and of higher levels of cysteine circulating in the bloodstream on the physiology of the sheep itself are unanswered and will remain so until a transgenic sheep producing elevated levels of cysteine is created. Although it is impossible to predict all of the consequences of such an achievement, some of the likely outcomes are considered in the discussion to follow, and are related to the metabolism of sulphur in mammalian cells, with special consideration given to the metabolism of sulphur in the sheep and the effect of rumen physiology upon this. Subsequently, there will be a discussion of the methods available which would improve the chance of success of the current transgenesis programme. This will cover the use of alternative cysteine biosynthesis genes, an examination of the structure and function of the linked-transgene pair and also of the mode of presentation of the transgene sequences to the sheep genome. A brief description of the new methods available for the production of transgenic sheep will be given, but the related technical details are beyond the scope of this thesis.

6.2 Sulphur metabolism in mammals.

Sulphur and the sulphur-containing amino acids are known to play an important role in supporting the rapid growth phases of young animals (e.g. as seen by improved nitrogen retention in lambs fed elemental sulphur, sodium

sulphate or DL-methionine whilst on a sulphur deficient ration; Starks et al.,1954) and also in the development and maintenance of structures, particularly in the keratinous tissues such as skin, feathers and wool, of adult animals. The improved rate of wool growth upon abomasal infusion of cysteine is an example of the latter (see section 1.3 and Reis and Schinckel, 1963, 1964). As the aim of the ongoing project is to provide an alternative means of de novo cysteine synthesis to mammalian cells (i.e. to the epithelial cells lining the sheep rumen) the remainder of this section will involve a consideration of the possible ramifications of the operation of this pathway in terms of the normal metabolic fates of intracellular cysteine. In particular, the likely consequences of production of quantities of cysteine in excess of that produced normally via transulphuration will be discussed. This is of paramount importance, considering the various inborn errors of sulphur metabolism which have been described in mammals, for example, cystinosis, involving an accumulation of cystine, homocystinurias I-IV, and sulphituria due to an inability to convert sulphite to sulphate prior to its excretion. Each of these demonstrate the result of interference to the normal mechanisms of sulphur metabolism.

In mammalian cells, the sulphur-containing amino acids, cysteine and methionine are involved in a multitude of metabolic processes and are related through the transulphuration pathway in which methionine is converted to cysteine (see section 1.4). With respect to cysteine, some of these processes are directly responsible for the metabolism of cysteine itself and help to maintain intracellular cysteine at a comparably low level, about 30-250 μM (Weinstein et al., 1988). Other processes are related to the synthesis and degradation of different sulphur-containing molecules, particularly proteins, glutathione, pantetheine and coenzyme A.

The fact that cysteine is only found at a low level in cells is not surprising considering its reactive nature. Glutathione on the other hand, a tri-peptide which contains cysteine, is found at intracellular levels of 0.5-10 mM in animal tissues (Finkelstein et al., 1982) and in human red blood cells at 2.3-2.8 mM, compared with cysteine levels of only 8-10 μM in these cells (Fahey et al., 1981) and is thought to serve partly as a cysteine-storage molecule. In fact, in vivo studies of glutathione turnover clearly suggest that its roles are to mediate inter-organ

cysteine transport and also to regulate and maintain the cellular pool of free cysteine (Weinstein et al., 1988). The currency of intracellular cysteine is therefore actually glutathione. This molecule, whose synthesis is usually limited by the availability of cysteine, provides a pool of cysteine which is in a metabolically active form, yet is non-toxic and can be easily transported (Griffith and Meister, 1979). Synthesized predominantly in the liver and kidney, glutathione is released into the plasma but found only at a low concentration there, approximately 25 μM (Anderson and Meister, 1980). Cysteine is released from glutathione by the action of enzymes localised on the external plasma membrane surface of cells and then taken up by these cells. In combination with the cysteine derived from methionine by transulphuration, this forms the intracellular pool of free cysteine. Of this, greater than 90 % is maintained in its reduced form by transhydrogenation with intracellular glutathione, which itself is kept in a reduced state by glutathione reductase (Griffith, 1987).

Importantly, though cysteine is used in glutathione synthesis in the cell, there is no intracellular mechanism by which glutathione may be converted to cysteine. Hence, some of the cysteine synthesized intracellularly in the rumen epithelium would first be converted to glutathione, then move from the cell in this form into the plasma, from whence it would be taken for use by other cells. In this respect, increased levels of intracellular cysteine may initially result in increased plasma levels of glutathione. If the appropriate membrane transport system for glutathione exists, cells such as those of the outer root sheath of the wool follicle, in closest proximity to the blood supply, would then have access to more cysteine via glutathione. One factor which would limit the entry of cysteine into this pathway is that the first of the two steps of glutathione synthesis is feedback inhibited by glutathione, which competes with glutamate at the active site of the enzyme γ-glutamyl-cysteine synthetase (Richman and Meister, 1975). It is expected that the K_i of glutathione as a competitive inhibitor, reported to be 2.3 mM for this enzyme in rat kidney (Richman and Meister, 1975) would set the limit for the amount of glutathione that could be synthesized in a cell, depending upon the rate of movement of glutathione from the cell.

Cysteine which is in excess of the capacity of this system could be used for intracellular synthesis of proteins, pantetheine or coenzyme A. Any remaining

cysteine not utilised would possibly create an imbalance in the reducing power in the cell. For example, auto-oxidation to cystine which occurs readily at neutral pH would produce H_2O_2 (Anderson and Meister, 1987). Formation of large quantities of peroxide either inside or outside the cell may be deleterious. However, as would be expected of cysteine in glutathione, one useful function in which excess free cysteine may be involved is the capture and detoxification of free radicals found in the cell, such as those of oxygen and carbon (Packer, 1974). This may in fact serve to counteract the formation of peroxide. High levels of free cysteine may interfere with the transport and function of certain metal ions and thereby disturb the normal metabolic functioning of the cell (i.e. many enzyme molecules require the presence of metal ions for conformation and catalysis at the active site, and the regulation of expression of some genes is effected by the involvement of metal ions).

Another property of cysteine is its ability to react with ketones (e.g. pyruvate) and reactive aldehydes such as glyoxylate and pyridoxal 5-phosphate. Where cysteine is known to inhibit pyridoxal phosphate-requiring enzymes, it is suspected that either a cyclic adduct is formed at the active site which then prevents further catalysis (Cooper, 1983) or that interaction of cysteine with free pyridoxal phosphate may deplete cells of that essential coenzyme and thus inhibit enzyme activity (Griffith, 1987). The toxicity of high concentrations of cysteine to the enzymes of the central nervous system is also possibly related to the interaction of cysteine with the coenzymes utilised by these enzymes. Likewise, the formation of S-adducts with essential carbonyls is thought to play a role in the accelerated atteroschlerosis observed in some disorders of sulphur amino acid metabolism (Cooper, 1983). Hence problems affecting normal function at the level of the single cell and the whole organism, may possibly arise as a result of the production of large quantities of cysteine intracellularly.

The preceding discussion suggests that synthesis of excessive amounts of cysteine intracellularly would not be desirable. Further, experimental determination of the optimum level of cysteine synthesis in mammalian cells may be a prerequisite to success in this sheep transgenesis programme. To this end, the results of preliminary short-term (one hour) experiments aimed to test cysteine synthesis in situ in transfected CHO cells indicate that they can support elevated

levels of cysteine synthesis and show no sign of deterioration during the assay period (A.V. Sivaprasad, unpublished data). Extension of the assay period in future experiments would provide better in vitro evidence of the ability of mammalian cells to endure constitutive cysteine synthesis mediated by an introduced pathway.

In any event, cysteine produced at high levels intracellularly would be expected to leave the cell of origin via a cysteine concentration gradient established there. This would also be dependent on the action of the alanine-serine-cysteine (ASC) amino acid transport system (Christensen et al., 1967; Shotwell et al., 1980 and Young et al., 1988) in maintenance of intracellular levels of cysteine. In this case, cysteine may not have an opportunity to exert the above-mentioned deleterious effects on the cell. This will be discussed further in relation to cysteine synthesis in the sheep rumen epithelium, in section 6.3.

Although there are many metabolic options for intracellular cysteine, the two main catabolites of this amino acid are sulphate, excreted in the urine, and taurine, found in tissues (Cooper, 1983). The origin of each of these end products involves the oxidation of cysteine to form cysteine sulphinate, a reaction believed to be the major catabolic pathway of cysteine in mammals, especially when cysteine availability is high. Synthesis of increased quantities of cysteine will lead to the formation of more of these catabolites. Taurine, originally thought to be relatively inert in mammalian tissues, is found free in millimolar concentrations in tissues that are excitable (e.g. 50-80 mM in the retinal pigmented epithelium and photoreceptor cell layer) and is in abundance in membranes that generate oxidants. Most lines of evidence suggest that the ultimate role for taurine lies in the stabilisation of membranes (Wright et al., 1986). The fidelity of sulphate metabolism is equally critical. For example, malfunction or deficiency of the enzyme sulphite oxidase which converts sulphite to sulphate results in severe neurological disorders, as observed in the case of sulphituria in humans. Considering the normally high rate of excretion of sulphate measured in humans (20-30 mmoles/day) the sulphur imbalance created is readily understood. In sheep, the urinary loss of sulphate is over the range of about 1-30 mmoles/day, depending on the level of sulphur in the diet (Kennedy and Milligan, 1978) and an equally important sulphur balance exists (see section 6.3). In

transgenic sheep producing elevated levels of cysteine, sulphite may accumulate if the capacity of sulphite oxidase is exceeded, and a similar disorder may arise.

It is obvious that many changes to normal cell metabolism could result from synthesis of extra cysteine de novo in the cells of the sheep rumen epithelium. However, the key to the production of increased levels of intracellular cysteine lies in the availability of the required substrates. Cysteine biosynthesis gene products expressed in the basal layers of the sheep rumen epithelium should have access to serine, acetyl CoA and sulphide. It is clear from the gene expression studies completed to date that sufficient enzyme can be synthesized in a cell to support the required level of cysteine synthesis. The amount of cysteine made by these cells per day will be limited by the intracellular concentration of these three substrates. Considering the amount of sulphide moving through the rumen epithelium each day and the K_m of the OAS sulphydrylase-A enzyme for this substrate (see section 6.4.1) sulphide availability is not expected to be a limiting factor in this situation. Instead, it seems likely that the normal intracellular concentrations of serine and acetyl CoA will set the limits for cysteine synthesis by the introduced pathway.

6.3 Sulphur metabolism in sheep : rumen physiology and the sulphur balance.

The metabolism of sulphur in sheep has been intensely researched since the discovery that the rate of wool growth is limited by the supply of sulphur amino acids to the follicle. Many attempts have been made to quantitate the movements of sulphur in the body of the sheep (Bray, 1969a, 1969b; Bray and Till, 1975; Kennedy et al., 1976; Kennedy and Milligan, 1978; Doyle and Moir, 1979). The result of this work suggests the obvious; that many parameters affect sulphur metabolism in the sheep and that maintenance of a balance reflects a complex interplay of these. In addition, the loss of significant quantities of sulphur from the sheep is supported by a number of the experimental findings described.

A subsidiary aim of the work described in this thesis is capture of the sulphur normally lost from the sheep, up to about 1.0 g of sulphur/day as sulphate in the urine. Conversion of this to a useable form, namely cysteine,

could result in greater sulphur retention in the sheep if the sheep did not then mobilise sulphur available in other compartments for excretion in order to maintain its sulphur balance. The results of cysteine infusion experiments suggest that this is not the case, and in fact more sulphur is ultimately retained by the sheep as wool keratin protein (Reis, 1989). The use of cysteine sulphur in the synthesis of greater quantities of wool could be considered the way in which the sheep maintains a balance of sulphur in the body, as this ultimately represents a loss of sulphur from the sheep, albeit in a more useful form.

In relation to the maintenance of this balance, the movement of sulphur throughout the intracellular and extracellular compartments is subject to a complex array of factors. Superimposed upon the usual inter-relationships between the use of exogenous or endogenous sulphur to provide for the bodily sulphur requirements are the sulphur requirements of an active microbial population resident in the sheep rumen. Apart from the normal dissimilatory action these microorganisms exert on dietary sulphur compounds in the rumen, leading to the production of sulphide (see section 1.4), sulphur is also incorporated into microbial protein and returns to the sheep following digestion in the abomasum (see figure 1.1). In addition, these microorganisms have access to the sulphur from sulphate reabsorbed by the kidney. Hence, assuming that the production of higher levels of cysteine in the sheep is possible and that this will probably affect the level of sulphate produced, the response of the sheep to this will also depend upon the response of the microbe population. Although indirect benefit to these microbes is quite likely as a result of increased levels of cysteine in the rumen epithelium, direct benefit to the rumen microorganisms by the use of free cysteine is thought unlikely and the reasons are discussed below.

If the expected movement of cysteine from a cell following a concentration gradient is considered in relation to cysteine biosynthesis in the rumen epithelium, it is reasonable to assume that excess cysteine would leave these cells and enter the extracellular compartments, including the bloodstream. The rumen epithelium, typical of keratinised epithelia, is comprised of four main layers (Steven and Marshall, 1970). These are supported by a well vascularised connective tissue sub-stratum. The outermost layer of this epithelium, the stratum corneum, is highly keratinised and serves as a protective barrier between the soft tissues of the

sheep and the fermentative processes occurring in the lumen of the rumen. Evidently, there is negligible absorption of cysteine and methionine from the rumen through this epithelium (Leibholz, 1971). By the same token, it is expected that apart from that used in rumen keratin protein synthesis, the cysteine synthesized in cells of the more metabolically active inner layers (i.e. the stratum basale, spinosum and granulosum) would preferentially enter the bloodstream and surrounding tissue rather than traverse the stratum corneum and enter the lumen. Hence, assuming that losses to the lumen would be minimal and to ensure maximum transfer of synthesized cysteine to the bloodstream, the ideal site of expression of the cysteine biosynthesis genes is in the stratum basale, the epithelial layer in closest proximity to the blood vessels and furthest from the lumen. (The data suggests that the amount of sulphide passing from the lumen through this epithelium will be sufficient to support cysteine synthesis targeted for the basal layer; see section 1.7.) In this case, it is estimated that the rumen microorganisms will only derive metabolic gain from the cysteine incorporated into rumen proteins synthesized in the stratum granulosum and stratum corneum, and not from free cysteine synthesized in the stratum basale. If intracellular cysteine surplus to the normal cellular level is produced in the stratum basale, then it will be utilised in one of the variety of ways described above (section 6.2). Ultimately, cysteine (possibly free, but probably in glutathione) will enter the plasma and be transported throughout the body. In the sheep, it has been shown that free cysteine reacts with the thiol groups of plasma proteins in vivo to form mixed disulphides which may play a role in cysteine transport to the follicle (Downes, 1961). Importantly, even if this cysteine is consumed by other cells prior to reaching the wool follicle, the pool of available cysteine would have been increased, and will provide for improved wool growth.

One of the known consequences of increased plasma levels of cyst(e)ine in sheep, originally observed in infusion experiments, is the alteration to the levels of other circulating amino acids (Reis et al., 1973b). Sheep were given abomasal infusions of L-methionine and L-cystine and concomitant with an increase in wool growth rate was a decrease in total plasma amino acids in each case. Assuming the use of these in wool synthesis, this is not necessarily surprising, but of particular interest is the considerable reduction in the level of

serine observed when either methionine or cystine were infused. A reduction in the plasma concentration of serine from 17.8 to 10.4 μmoles/100 ml of plasma (a 42 % reduction) was seen upon daily infusion of 2.5 g of L-methionine, and from 18.9 to 11.9 μmoles/100 ml of plasma (a 37 % reduction) upon daily infusion of 2.0 g of L-cystine (Reis et al.,1973). Only slightly greater reductions in the plasma serine concentration were seen on infusion of larger quantities of L-methionine or L-cystine, suggesting that the limits of this phenomenon had nearly been reached at the lower infusion rates mentioned. Considerable reductions in the plasma concentrations of valine, leucine, isoleucine and glycine were also measured. In addition, an increase in the level of taurine was observed in each case.

As a result of cysteine synthesis in the rumen, plasma serine levels may drop for two reasons. Initially, serine consumed in the first step of cysteine synthesis may cause a reduction in plasma serine concentration. Secondly, higher plasma levels of cysteine may cause a further reduction of plasma serine via the mechanism operating in the infusion experiments. This will in turn affect the cysteine synthesis reaction, hence these inter-related effects should reach an equilibrium. At this stage it is impossible to predict whether this will be a useful inbuilt limitation of the system.

6.4 Improvements to the transgenesis regime.

In the past decade, genetic engineering and improvement in the techniques associated with the culturing of mammalian cells and embryos, including their micromanipulation, has enabled vast accomplishments in many different spheres of medical and veterinary science, and in agriculture. One facet of the technology which is yet to realise its full potential is the ability to alter individual genotypes (and hence phenotypes) through transgenesis. The research programme described in the preceding chapters, the aim of which is to achieve expression of the bacterial cysteine biosynthesis pathway in the sheep (rumen) and ultimately improve wool growth rate, is an example of the application of this technology. To reach this goal, improvements to the current programme are necessary. These include optimisation of expression of the cysteine biosynthesis genes, testing of alternative genes and the achievement of specificity of transgene expression.

6.4.1 Optimisation of expression of cysteine biosynthesis transgenes.

In order to achieve tissue-specific coexpression of cysteine biosynthesis genes in the sheep, expression of the genes in the linked-gene constructs requires optimisation (see Discussion section 5.6.2). This could be realised primarily through manipulation of the sequences flanking the coding regions of each of the genes to improve transcription and translation. At the level of transcription, absolute control through an inducible promoter is not sought in this case, except that expression of the transgenes should ideally be restricted to the adult animal in order to eliminate problems which could arise due to a higher availability of cysteine in the developing foetus and neonate.

In the first instance, replacement of the RSVLTR promoter with a rumen-specific promoter is ideal. In this laboratory, a number of genes expressed in the rumen epithelium and other epithelia have been isolated (L.A. Whitbread and C.J. McLaughlan, unpublished data). Through in situ hybridisation using cRNA probes, these have been shown to be expressed in particular cell layers of the rumen epithelium. Northern analysis suggests that they are usually expressed at a moderate to high level in the adult rumen epithelium. One of the genes is expressed predominantly in the stratum basale, the preferred site of expression considering its proximity to the capillary network and according to the information available regarding sulphide flux through this tissue. Currently, definition of the sequences required to direct the tissue-specific expression of these genes is being sought. This will involve testing of the promoter elements in a number of expression systems and ultimately in transgenic sheep. Due to the long gestation period of sheep, it is hoped that short-term experiments involving expression of test constructs in a primary adult rumen cell line or in an in vitro nuclear extract prepared from adult rumen cells will be possible and minimise testing by sheep transgenesis.

Other promoters responsible for directing the expression of genes involved in intracellular "house-keeping" functions are being considered for use in this project. One such promoter being tested currently is the mouse phosphoglycerate kinase promoter (Chaker et al., 1987). Promoters such as this

have an advantage in that they may allow a low to moderate level of constitutive expression of the cysteine synthesis genes. Considering the consequences of overproduction of intracellular cysteine, a less active promoter may be desirable. In the case of the promoter of a house-keeping gene which functions constitutively in all cells, substrate availability would limit the production of cysteine (via the second step reaction) to the rumen epithelium. However, the first step reaction, production of OAS, may proceed in all tissues and unnecessarily utilise serine and acetyl CoA. Although OAS produced in vitro is slowly converted back to serine (A.V. Sivaprasad, unpublished observation) and the cysE transgenic mice suffered no ill-effect through expression of active SAT enzyme in many mesodermal tissues, the likely consequences of non-specific expression and SAT activity throughout vital organs such as the brain are not known.

Also related to transcription is the use of appropriate 3' processing and termination signals to maximise mRNA stability and minimise possible transcriptional interference, where expression of a gene in a transgene locus or a linked-gene construct is inhibited by expression of an upstream gene. Whilst the hGH 3' end used in the current constructs confers a high degree of stability on the resulting mRNA, transcriptional interference (Proudfoot, 1986; Kadesch and Berg, 1986; Platt, 1986) is of particular concern in coexpression of genes in the current linked-gene constructs, where the hGH polyadenylation signal of the 5' gene and the RSVLTR promoter cap site of the 3' gene are only separated by 1.8 kb. As outlined in section 5.6.1, the UMS polyadenylation/termination sequence is being compared with the hGH sequence in a new series of linked-gene constructs. This 3' end sequence has been reported to facilitate transcription termination, in addition to polyadenylation and 3' processing (McGeady et al., 1986, Clouston et al., 1990), and could minimise any transcriptional interference occurring during expression of genes in the linked-gene construct. When the 6.4 kb linked-transgene fragment is integrated into the host genome in multi-copy arrays, the UMS may also provide a degree of copy number dependence, depending on the constraints imposed by the host genome at a particular location.

An important aspect related to regulation of the transcription of genes in a transgene locus is the influence of the host genome. As previously discussed in

the literature (Palmiter and Brinster, 1986; Al-Shawi et al.,1990) and clear from the linked-gene cell line studies and expression levels in different cysE transgenic mouse lines, the influence of the host genome on expression of foreign genes can be considerable. Even if a tissue-specific promoter is available, the fidelity of expression of a randomly-integrated transgene may be severely affected if it is not isolated from chromosome position effects. A higher level of control of gene expression clearly exists and has been lacking in most of the transgenesis performed to date. Recently, a new DNA element known as the locus activation region (LAR) or dominant control region (DCR) has been identified (Forrester et al., 1987; Grosveld et al., 1987) and demonstrated to facilitate tissue-specific, copy number-dependent expression of transgenes irrespective of chromosomal location (Greaves et al., 1989; Bonifer et al.., 1990). Although a number of examples have been documented, it is not yet certain that LAR's exist for all genes. If this proved to be the case, isolation and use of LAR's from the rumen genes under investigation is an option which would provide a great advantage to the current transgenesis programme. Another DNA element which might be of similar use in transgenesis is the scaffold attachment/association region (SAR; also known as the matrix attachment/association region or MAR; Jarman and Higgs, 1988; Levy-Wilson and Fortier, 1989; Blasquez et al., 1989). Like the LAR, this has the ability to eliminate the effects that host genomic sequences flanking the transgene locus have on expression (Phi-Van et al., 1990).

At the level of translation, improvements could be made to each of the bacterial cysteine synthesis gene coding regions with respect to codon usage and the effect that this has on elongation during polypeptide synthesis. This could be achieved through mutagenesis of codons which consist of a sequence less commonly used in mammalian translation. Likewise, the efficiency of initiation of the cysE gene translation could be improved by removal of extraneous bacterial sequences and mutagenesis of the 5'-untranslated region to the optimal sequence reported by Kozak (see section 3.5.1). Optimisation of the sequences forming the translation initiation environment of the cysM coding region has already been achieved.

The use of alternative cysteine synthesis genes in order to optimise expression of a cysteine biosynthesis pathway in sheep is also under

consideration. In this regard, an <u>E. coli</u> cysE mutant which is less sensitive to feedback inhibition by the end product cysteine has been discussed previously (see section 5.6.1). An additional option being investigated at present is the use of a gene from <u>S. typhimurium</u>, cysK, that encodes the OAS sulphydrylase-A enzyme, which has a higher affinity for each of its substrates, OAS ($K_m = 0.15$ mM) and sulphide ($K_m = 0.066$ mM). The corresponding affinities of the cysM gene product, OAS sulphydrylase-B, for these substrates are 7.0 mM and 3.6 mM respectively (Kredich, 1987) about 50-fold lower in each case. According to these values, the combined difference in potential catalytic turnover of the substrates via these OAS sulphydrylase enzymes could be 2500-fold. To date, the cysK gene (a gift of Dr. N. Kredich) has been cloned into plasmid vectors and active enzyme has been expressed in cell culture (A.V. Sivaprasad, unpublished data). In addition, the cysK gene has been cloned in linked-gene constructs and successfully coexpressed with the cysE gene in permanent cell cultures (A.V. Sivaprasad, unpublished data) to reproduce the cysteine synthesis capacity first measured in extracts made from the pLT-EM cell lines. These initial experiments indicate high levels of OAS sulphydrylase-A activity and cysteine synthesis. More rigorous comparisons of these results with the OAS sulphydrylase-B (cysM gene product) activity and cysteine synthesis capacity of the previously cloned cell lines is currently in progress. However, if this gene proves to encode an enzyme product which has a greater activity than the cysM gene product in mammalian systems, then it will be an ideal replacement for the cysM gene in the sheep transgenesis programme.

6.4.2 Site-specific transgene integration.

Properly coordinated gene expression is critical in the case of linked-gene DNA constructions. Transgene fragments generally become arranged into multicopy tandem arrays in the host genome and the proximity of neighbouring genes in this situation (as in the cell cultures studied) and their effect upon each other must be considered. Whilst the effects of the host genome upon transgene expression cannot be overestimated, the damage which may be done to the development and growth of an organism because of insertional inactivation of genes by random integration of transgene sequences is undesirable. The use of

the embryonic stem-cell / homologous recombination approach to transgenesis should eliminate this problem. Ideally, this should be complemented by the added specificity afforded by such elements as locus activation regions. Unfortunately, as this technology is in its infancy, the immediate future of this transgenesis programme relies upon optimisation of the current working parameters, including those described above related to expression of the cysteine synthesis genes, and to the conventional methods of pronuclear microinjection and embryo culture methods established for producing transgenic sheep (Simons et al., 1988; Walker et al., 1990). Other than this, the yield of transgenic animals carrying the linked-transgene construct may only be improved by application of embryo-splitting techniques following the use of PCR to detect integrated transgenes as early as possible in development of the embryo. The use of these techniques is currently under investigation in this laboratory.

6.5 Testing for cysteine synthesis and improved wool growth.

Upon production of transgenic sheep expressing each of the cysteine biosynthesis genes in the rumen epithelium, analysis will necessarily begin with determination of the plasma levels of cysteine and/or glutathione. This will be supported by in vitro analysis of the cysteine synthesis capacity of primary cell cultures derived from adult rumen tissue explants. In addition, measurements of wool growth rate will be made via analysis of incorporation of ^{35}S into wool follicle proteins following pulse labelling of skin sections with tracer doses of L-^{35}S-cystine according to the method described by Downes (1965). Other features related to the general physiology of the sheep and to rumen physiology in particular will be carefully monitored (e.g. the levels of all amino acids in the plasma), so that any problems associated with greater sulphur retention may be rapidly detected.

In recent cysteine infusion experiments, many changes have been documented at the RNA and protein level, yet in these and past experiments, the ultimate consequences for fibre properties other than those of dimension have not been determined. To complement the physiological analyses, other tests related to the production of wool fibres will be performed, including analysis of the types

and levels of the IF and IFAP keratin proteins produced and, specifically, the physical properties of the resulting fibres. Thus, while production of transgenic sheep producing cysteine <u>de novo</u> may be of immediate commercial use, analysis of these sheep will provide further insight into the interaction of keratin proteins in the wool fibre and indicate the possible ways in which future genetic manipulation may be used to optimise wool growth and improve characteristics of the fleece.

Bibliography

Allden, W.G. (1979). in "Physiological and Environmental Limitations To Wool Growth." (Black, J.L. and Reis, P.J., eds.) University of New England Publishing Unit. pp.61-78.

Al-shawi, R. Kinnaird, J., Burke, J. and Bishop, J.O. (1990). Mol. Cell. Biol. 10, 1192-1198.

Anderson, M.E. and Meister, A. (1980). J. Biol. Chem. 255, 9530-9533.

Anderson, M.E. and Meister, A. (1987). Methods in Enzymology 143, 313-325.

Aota, S-i., Gojobori, T., Ishibashi, F., Maruyama, T. and Ikemura, T. (1988). Nucleic Acids Res. 16 (Suppl.) r315-402.

Auerswald, E.-A., Ludwig, G. and Schaller, H. (1980). Cold Spring Harbour Symp. Quant. Biol. 45, 107-113.

Bawden, C.S., Cam, G.R., D'Andrea, R.J., Kuczek, E.S., Powell, B.C., Rogers, G.E. and Sivaprasad, A.V. (1987). in "Merino Improvement Programs in Australia." Proceedings of a National Symposium, Leura N.S.W., 1987. (McGuirk, B.J. ed.) Australian Wool Corporation Technical Publication, pp.511-518.

Bird, P.R. (1972). Aust. J. Biol. Sci. 25, 185-193.

Birnboim, H.C. and Doly, J. (1979). Nucleic Acids Res. 7, 1513-1523.

Black, J.L. and Reis, P.J., eds. (1979)."Physiological and Environmental Limitations To Wool Growth." University of New England Publishing Unit, Armidale.

Blasquez, V.C., Xu, M., Moses, S.C. and Garrard, W.T. (1989). J. Biol. Chem. 264, 21183-21189.

Bonifer, C., Vidal, M., Grosveld, F. and Sippel, A.E. (1990). EMBO J. 9, 2843-2848.

Bray, A.C. (1969b). Aust. J. Agric. Res. 20, 725-737.

Bray, A.C. (1969a). Aust. J. Agric. Res. 20, 739-748.

Bray, A.C. and Till, A.R. (1975) in "Digestion and Metabolism in the Ruminant." (McDonald, I.W. and Warner, A.C.I. eds.). University of New England Publishing Unit. pp.243-260.

Brinster, R.L., Chen, H.Y., Trumbauer, M., Senear, A.W., Warren, R. and Palmiter, R.D. (1981). Cell 27, 223-231.

Brooker, J.D., Lockington, R.A., Attwood, G.T., Langridge, P., Nield, J.K. and Langridge, U. (1989) in "The Biology of Wool and Hair." (Rogers, G.E. et.al., eds.) Chapman and Hall. pp.425-440.

Brown, P.O., Bowerman, B., Varmus, H.E. and Bishop, J.M. (1987). Cell 49, 347-356.

Burnette, W.N. (1981). Anal. Biochem. 112, 195-203.

Burnham, B.F. and Lascelles, J. (1963). Biochem. J. 87, 462-472.

Byrne, C.R., Wilson, B.W. and Ward, K.A. (1987). Aust. J. Biol. Sci. 40, 459-468.

Capecchi, M.R. (1989). Science 244, 1288-1292.

Casadaban, M.J. and Cohen, S.N. (1980). J. Mol. Biol. 138, 179-207.

Cavener, D.R. (1987). Nucleic Acids Res. 15, 1353-1361.

Chaker, N.A., Poppo, H.B. and McBurney, M.W. (1987). Gene 60, 65-74.

Checchi, I., de Richardi (1961). Bull. Soc. Ital. Farm. Ospitial 7, p165.

Cheng, K.-J., McCowan, R.P. and Costerton, J.W. (1979). Am. J. Clin. Nutr. 32, 139-148.

Cherest, H., Eichler, F. and de Robichon-Szulmajster, H. (1969). J. Bacteriol. 97, 328-336.

Chomczynski, P. and Sacchi, N. (1987). Anal. Biochem. 162, 156-159.

Christensen, H.N., Liang, M. and Archer, E.G. (1967). J. Biol. Chem. 242, 5237-5246.

Chu, G., Hayakawa, H. and Berg, P. (1987). Nucleic Acids Res. 15, 1311-1326.

Clemens, M.J. (1984). in "Transcription and Translation : a practical approach." (Hames, B.D. and Higgins, S.J. eds.) IRL Press Ltd. pp.231-270.

Clouston, W.M., Lloyd, C.J. and Richards, R.I. (1990). J. Molec. Endocrinol. (in press).

Cooper, A.J.L. (1983). Ann. Rev. Biochem. 52, 187-222.

Cooper, J.D.H. and Turnell, D.C. (1987). Methods in Enzymology 143, 141-148

Crewther, W.G. (1976). Proc. Fifth Int. Wool Text. Res. Conf., Aachen 1975. 2, 1-103.

Dagert, M. and Ehrlich, S.D. (1979). Gene 6, 23-28.

D'Andrea, R.J., Sivaprasad, A.V., Bawden, C.S., Kuczek, E.S., Whitbread, L.A. and Rogers, G.E. (1989). in "The Biology of Wool and Hair." (Rogers, G.E. et.al., eds.) Chapman and Hall. pp.447-464.

D'Andrea, R.J., Surdin-Kerjan, Y., Pure, G. and Cherest, H. (1987). Mol. Gen. Genet. 207, 165-170.

Davis, L.G., Dibner, M.D. and Battey, J.F. (1986). "Basic Methods in Molecular Biology. A Laboratory Manual." Elsevier Science Publishing Co. Inc..

de Robichon-Szulmajster, H. and Cherest, H. (1967). Biochim. Biophys. Res. Comm. 28, 256-262.

Denk, D. and Bock, A. (1987). J. Gen. Microbiol. 133, 515-525.

Devereux, J., Haeberli, P. and Smithies, O. (1984). Nucleic Acids Res. 12, 387-395.

Downes, A.M. (1961). Aust. J. Biol. Sci. 14, 427-439.

Downes, A.M. (1965). in "Biology of the Skin and Hair Growth." (Lyne, A.G. and Short, B.F. eds.) Angus and Robertson Ltd., Sydney. pp.345-364.

Downes, A.M., Sharry, L.F. and Till, A.R. (1964). Aust. J. Biol. Sci. 17, 945-959.

Downes, A.M., Reis, P.J., Sharry, L.F. and Tunks, D.A. (1970). Aust. J. Biol. Sci. 23, 1077-1088.

Doyle, P.T. and Moir, R.J. (1979). Aust. J. Biol. Sci. 32, 65-76.

Fahey, R.C., Newton, G.L., Dorian, R. and Kosower, E.M. (1981). Anal. Biochem. 111, 357-365.

Faichney, G.J. and Black, J.L. (1979) in "Physiological and Environmental Limitations To Wool Growth." (Black, J.L. and Reis, P.J., eds.) University of New England Publishing Unit. pp.179-192.

Feenstra, A., Fewell, J., Lueders, K. and Kuff, E. (1986). Nucleic Acids Res. 14, 4343-4352.

Finkelstein, J.D., Kyle, W.D. Harris, B.J. and Martin, J.J. (1982). J. Nutr. 112, 1011-1018.

Forrester, W.C., Novak, U., Gelinas, R. and Groudine, M. (1989). Proc. Natl. Acad. Sci. USA 86, 5439-5443.

Forster, A.C., McInnes, J.L., Skingle, D.C. and Symons, R.H. (1985). Nucleic Acids Res. 13, 745-761.

Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. and Turner, D.H. (1986). Proc. Natl. Acad. Sci. USA 83, 9373-9377.

Frenkel, M.J., Gillespie, J.M. and Reis, P.J. (1974). Aust. J. Biol. Sci. 27, 31-38.

Frenkel, M.J., Gillespie, J.M. and Reis, P.J. (1975). Aust. J. Biol. Sci. 28, 331-338.

Freshney, R.I. (1987) in "Culture of Animal Cells. A Manual of Basic Technique." A.R. Liss, Inc., N.Y., pp.137-153.

Friedman, M., Krull, L.H. and Cavins, J.F. (1970). J. Biol. Chem. 254, 3868-3871.

Fujiwara, T. and Mizuuchi, K. (1988). Cell 54, 497-504.

Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M.A. (1987a). Nucleic Acids Res. 15, 3257-3273.

Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M.A. (1987b). Nucleic Acids Res. 15, 8693-8711.

Gillespie, J.M. and Broad, A. (1972). Aust. J. Biol. Sci. 25,139-145.

Gillespie, J.M., Broad, A. and Reis, P.J. (1969). Biochem. J. 112, 41-49.

Gillespie, J.M., Reis, P.J. and Schinckel, P.G. (1964). Aust. J. Biol. Sci. 17, 548-560.

Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982). Mol. Cell. Biol. 2, 1044-1051.

Gough, N.M., Metcalf, D., Gough, J., Grail, D. and Dunne, A.R. (1985). EMBO J. 4, 645-653.

Greaves, D.R., Wilson, F.D., Lang, G. and Kioussis, D. (1989). Cell 56, 979-986.

Greenstein, J.P. and Winitz, M. (1967) in "Chemistry of the Amino Acids." John Wiley and Sons Inc., p.1053 and p.2230.

Griffith, O.W. (1987). Methods in Enzymology 143, 366-376.

Griffith, O.W. and Meister, A. (1979). Proc. Natl. Acad. Sci. USA 76, 5606-5610.

Grosveld, F., van Assendelft, G.B., Greaves, D.R. and Kollias, G. (1987). Cell 51, 975-985.

Heard, J.-M., Herbomel, P., Ott, M.-O., Mottura-Rollier, A., Weiss, M. and Yaniv, M. (1987). Mol. Cell. Biol. 7, 2425-2434.

Henderickx, H. (1961). Archs Int. Physiol. Biochim. 69, 443.

Higgins, T.J.V., O'Brien, P.A., Spencer, D., Schroeder, H.E., Dove, H. and Freer, M. (1989) in "The Biology of Wool and Hair." (Rogers, G.E. et.al., eds.) Chapman and Hall. pp.441-445.

Hogan, B., Constantini, F. and Lacy, E. (1986). "Manipulating the Mouse Embryo." Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.

Hogan, J.P., Elliott, N.M. and Hughes, A.D. (1979) in "Physiological and Environmental Limitations To Wool Growth." (Black, J.L. and Reis, P.J., eds.) University of New England Publishing Unit. pp.43-59.

Hulanicka, M.D. and Kredich, N.M. (1976). Mol. Gen. Genet. 148, 143-148.

Hulanicka, M.D., Kredich, N.M. and Treiman, D.M. (1974). J. Biol. Chem. 249, 867-872.

Humphries, R.K., Ley, T., Turner, P., Moulton, A.D. and Nienhuis, A.W. (1982). Cell 30, 173-183.

Huovinen, J.A. and Gustafsson, B.E. (1967). Biochem. Biophys. Res. Comm. 136, 441-447.

Hynd, P.I. (1989) in "The Biology of Wool and Hair." Chapman and Hall. (Rogers, G.E. et.al., eds.) pp.169-184.

Ish-Horowitz, D and Burke, J.F. (1981). Nucleic Acids Res. 9, 2989-2998.

Jaenisch, R., (1988). Science 240, 1468-1474.

Jarman, A.P. and Higgs, D.R. (1988). EMBO J. 7, 3337-3344.

Jenkins, B.J. (1990). Honours Thesis, University of Adelaide, South Australia.

Jobling, S.A. and Gehrke, L. (1987). Nature 325, 622-625.

Jones, E.W. and Fink, G.R. (1982) in "The molecular biology of the yeast Saccharomyces : Metabolism and gene expression." (Strathern, J.N. et. al. eds.) Cold Spring Harbour Laboratory Press, N.Y.. pp.181-299.

Kadesch, T. and Berg, P. (1986). Mol. Cell. Biol. 6, 2593-2601.

Kennedy, P.M., Hogan, J.P., Lindsay, J.R. and Hogan, R.M. (1976). Aust. J. Biol. Sci. 29, 525-531.

Kennedy, P.M. and Milligan, L.P., (1978). British J. Nutrition 39, 65-83.

Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984). Cell 38, 731-736.

Kozak, M. (1981). Nucleic Acids Res. 9, 5233-5252.

Kozak, M. (1984). Nucleic Acids Res. 12, 857-871.

Kozak, M. (1987). Nucleic Acids Res. 15, 8125-8148.

Kredich, N.M. (1987) in "Molecular and Cellular Biology of E. coli and Salmonella." (Neidhardt, F.C. ed.) Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York pp. 1231-1240.

Kredich, N.M., Becker, M.A. and Tomkins, G.M. (1969). J. Biol. Chem. 244, 2428-2439.

Kredich, N.M. and Tomkins, G.M. (1966). J. Biol. Chem. 241, 4955-4965.

Krieg, P.A. and Melton, D.A. (1984). Nucleic Acids Res. 12, 7057-7070.

Krieg, P.A. and Melton, D.A. (1987). Methods in Enzymology 155, 397-415.

Laemmli, U.K. (1970). Nature 227, 680-685.

Layne, E. (1957). Methods in Enzymology 3, 447-454.

Lee, J.J. and Costlow, N.A. (1987). Methods in Enzymology 152, 633-648.

Leibholz, J. (1971). Aust. J. Agric. Res. 22, 639-645.

Levy-Wilson, B. and Fortier, C. (1989). J. Biol. Chem. 264, 21196-21204.

Lewis, D. (1954). Biochem. J. 56, 391-399.

Lopata, M., Cleveland, D. and Sollner-Webb, B. (1984). Nucleic Acids Res. 12, 5707-5717.

Luciw, P.A., Bishop, J.M., Varmus, H.E. and Capecchi, M.R. (1983). Cell 33, 705-716.

Marshall, R.C. and Gillespie, J.M. (1977). Aust. J. Biol. Sci. 30, 389-400.

Masselot, M. and de Robichon-Szulmajster, H. (1975). Mol. Gen. Genet. 139, 121-132.

McGeady, M.L., Wood, T.G., Maizel, J.V. and Vande Woude, G.F. (1986). DNA 5, 289-298.

McGuirk, B.J. ed. (1987). "Merino Improvement Programs in Australia." Proceedings of a National Symposium, Leura N.S.W., 1987. Australian Wool Corporation Technical Publication.

McKnight, S.L., Gavis, E.R., Kingsbury, R.C. and Axel, R. (1981). Cell 25, 385-398.

Mietz, J.A. and Kuff, E.L. (1990). Proc. Natl. Acad. Sci. USA 87, 2269-2273.

Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981). Cell 27, 279-288.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984). Nucleic Acids Res. 12, 7035-7056.

Messing, J. (1983). Methods in Enzymology 101, 20-78.

Monroe, R.S. and Kredich, N.M. (1988). J. Bacteriol. 170, 42-47.

Nordstrom, J.L., Hall, S.L. and Kessler, M.M. (1985). Proc. Natl. Acad. Sci. USA 82, 1094-1098.

Overbeek, P.A., Lai, S-P, Van Quill, K.R. and Westphal, H. (1986). Science 231, 1574-1577.

Packer, J.E. (1974) in "The Chemistry of the Thiol Group." (S. Patai ed.) Wiley, N.Y., p481.

Palmiter, R.D. and Brinster, R.L. (1986). Ann. Rev. Genet. 20, 465-499.

Pelham, H.R.B. and Jackson, R.J. (1976). Eur. J. Biochem. 67, 247-256.

Pfarr, D.S., Rieser, L.A., Woychik, R.P., Rottman, F.M., Rosenberg, M. and Reff, M.E. (1986). DNA 5, 115-122.

Phi-Van, L., von Kries, J.P., Ostertag, W. and Stratling, W.H. (1990). Mol. Cell. Biol. 10, 2302-2307.

Platt, T. (1986). Ann. Rev. Biochem. 55, 339-372.

Porra, R.J. and Jones, O.T.G. (1963). Biochem. J. 87, 181-185.

Powell, B.C. and Rogers, G.E. (1986) in "Biology of the Integument, 2, Vertebrates." (Bereiter-Hahn, J. et.al., eds.) Springer-Verlag, Berlin, pp.695-721.

Proudfoot, N.J. (1986). Nature 322, 562-565.

Pursel, V.G., Pinkert, C.A., Miller, K.F., Bolt, D.J., Campbell, R.G., Palmiter, R.D., Brinster, R.L. and Hammer, R.E. (1989). Science 244, 1281-1288.

Radcliffe, B.C. and Egan, A.R. (1974). Aust. J. Biol. Sci. 27, 465-471.

Reed, K.C. and Mann, D.A. (1985). Nucleic Acids Res. 13, 7207-7221.

Reis, P.J., (1967). Aust. J. Biol. Sci. 20, 809-825.

Reis, P.J. (1979) in "Physiological and Environmental Limitations To Wool Growth." (Black, J.L. and Reis, P.J., eds.) University of New England Publishing Unit. pp.223-242.

Reis, P.J. (1989) in "The Biology of Wool and Hair." Chapman and Hall. (Rogers, G.E. et.al., eds.) pp.185-203.

Reis, P.J. and Schinckel, P.G. (1963). Aust. J. Biol. Sci. 16, 218-230.

Reis, P.J. and Schinckel, P.G. (1964). Aust J. Biol. Sci. 17, 532-547.

Reis, P.J., Tunks, D.A. and Downes, A.M. (1973a). Aust. J. Biol. Sci. 26, 249-258.

Reis, P.J., Tunks, D.A. and Sharry, L.F. (1973b). Aust. J. Biol. Sci. 26, 635-644.

Richman, P.G. and Meister, A. (1975). J. Biol. Chem. 250, 1422-1426.

Rogers, G.E. (1990). Trends in Biotechnology 8, 6-11.

Rogers, G.E., Kuczek, E.S., MacKinnon, P.J., Presland, R.B., Fietz, M.J. (1989) in "The Biology of Wool and Hair." (Rogers, G.E. et.al., eds.) Chapman and Hall. pp 69-85.

Rogers, J.H. (1985). Int. Rev. Cytol. 93, 187-279.

Rothstein, S.J., Jorgensen, R.A., Yin, J.C.-P., Yong-di, Z., Johnson, R.C. and Reznikoff, W.S. (1980). Cold Spring Harbour Symp. Quant. Biol. 45, 99-105.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). Science 239, 487-491.

Sanger, F.C., Nicklen, S. and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Sangsoda, S., Cherest, H. and Surdin-Kerjan, Y. (1985). Mol. Gen. Genet. 200, 407-414.

Schlossmann, K. and Lynen, F. (1957). Biochem. Z. 328, 591-594.

Shih, C-C., Stoye, P.J. and Coffin, J.M. (1988). Cell 53, 531-537.

Shotwell, M.A., Jayme, D.W., Kilberg, M.S. and Oxender, D.L. (1981). J. Biol. Chem. 256, 5422-5427.

Simons, J.P., Wilmut, I., Clark, A.J., Archibald, A.L., Bishop, J.O. and Lathe, R. (1988). Biotechnology 6, 179-183.

Sleat, D.E., Hull, R, Turner, P.C. and Wilson, T.M.A. (1988). Eur. J. Biochem. 175, 75-86.

Sleigh, M.J. (1986). Anal. Biochem. 156, 251-256.

Soda, K. (1987). Methods in Enzymology 143, 453-459.

Southern, E.M. (1975). J. Mol. Biol. 98, 503-518.

Southern, P.J. and Berg, P. (1982). J. Mol. Appl. Genet. 1, 327-341.

Staden, R. (1984). Nucleic Acids Res. 12, 521-538.

Starks, P.B., Hale, W.H., Garrigus. U.S., Forbes, R.M. and James, M.F. (1954). J. Animal. Sci. 13, 249-257.

Steven, D.H. and Marshall, A.B. (1970) in "Physiology of Digestion and Metabolism in the Ruminant." (Phillipson, A.T. ed.) Oriel Press, Newcastle upon Tyne, pp.80-100.

Van Tol, R.G. and Van Vloten-Doting, L. (1979). Eur. J. Biochem. 93, 461-468.

Varmus, H.E. and Swanstrom, R. (1985) in "RNA Tumour Viruses." (Weiss, R. et. al., eds.) Cold Spring Harbour Laboratory Press, N.Y., pp. 75-134

Vize, P.D. (1987). Ph. D. Thesis, University of Adelaide, South Australia.

Vogelstein, B. and Gillespie, D. (1979). Proc. Natl. Acad. Sci. USA 76, 615-619.

Waechter, D.E. and Baserga, R. (1982). Proc. Natl. Acad. Sci. USA 79, 1106-1110.

Walker, S.K. (1990). Ph. D. Thesis, University of Adelaide, South Australia.

Walker, S.K., Heard, T.M., Verma, P.J., Rogers, G.E., Bawden, C.S., Sivaprasad, A.V., McLaughlin, K.J. and Seamark, R.F. (1990). Reprod. Fertil. Dev. 2, 633-640.

Wang, Y-H., Sczekan, S.R. and Theil, E.C. (1990). Nucleic Acids Res. 18, 4463-4468.

Weinstein, C.L., Haschemeyer, R.H. and Griffith, O.W. (1988). J. Biol. Chem. 263, 16568-16579.

Williams, A.J. (1979) in "Physiological and Environmental Limitations To Wool Growth." (Black, J.L. and Reis, P.J., eds.) University of New England Publishing Unit. pp.337-354.

Wood, T.G., McGeady, M.L., Baroudy, B.M., Blair, D.G. and Vande Woude, G.F. (1984). Proc. Natl. Acad. Sci. USA 81, 7817-7821.

Wright, C.E., Tallan, H.H. and Lin, Y.Y. (1986). Ann. Rev. Biochem. 55, 427-453.

Yamagata, S. (1976). J. Biochem. 80, 787-797.

Yamagata, S. (1984). J. Biochem. 96, 1511-1523.

Yamagata, S. (1987). Methods in Enzymology 143, 478-483.

Yamamoto, T., de Crombrugghe, B. and Pastan, I. (1980). Cell 22, 787-797.

Young, J.D., Mason, D.K. and Fincham, D.A. (1988). J. Biol. Chem. 263, 140-143.

Young, M.R., Tolbert, B and Utter, M.F. (1969). Methods in Enzymology 13, 250-258.

Zuker, M. and Stiegler, P. (1981). Nucleic Acids Res. 9, 133-148.